

VERY HIGH-FREQUENCY TECHNIQUES

COMPILED BY THE STAFF OF THE
RADIO RESEARCH LABORATORY
HARVARD UNIVERSITY

*Office of Scientific Research and Development
National Defense Research Committee
Division of Radio Coordination*

UNDER THE EDITORIAL DIRECTION OF
Herbert J. Reich, Editor
Louise S. McDowell, Asst. Editor
Andrew Alford, Robert R. Buss, John F. Byrne
Arthur Dorne, William G. Dow, John D. Kraus
Joseph M. Pettit, Edwin A. Yunker

Volume I

VERY HIGH-FREQUENCY TECHNIQUES

VI

PREFACE

The characteristics of the circuits and techniques treated in this book were of necessity chosen primarily to meet the needs of the counter-measures program. Among the requirements were broad-band modulation and amplification, wide frequency range, and simplicity of tuning. This book does not pretend to be either a textbook or a comprehensive treatise covering all aspects of the u-h-f field. As a handbook or reference work, however, it should enable research physicists and engineers to save considerable time and effort in the solution of problems in this field. It will undoubtedly also serve as a valuable source of information to authors of textbooks. The material is sufficiently fundamental and new, and the range of subjects covered so great, that every worker in the v-h-f field will find it least some part of the material of interest.

The successful completion of this volume is due in no small measure to the efforts of Dr. F. E. Terman, who organized the material and guided its preparation.

HERBERT J. REICH

CAMBRIDGE, MASS.,
April, 1946.

FOREWORD

The Radio Research Laboratory was a large central laboratory established by the Office of Scientific Research and Development for developing countermeasures against enemy radar. It was operated by Harvard University and came under Division 15 of the National Defense Research Committee of which Dr. C. G. Suits was Division Chief.

The Radio Research Laboratory was unique in being the only large temporary wartime OSRD laboratory concerned in a major way with the extension of continuous-wave techniques to the high radio frequencies used in radar work. Such techniques were the basis of most prewar activity in radio below 100 Mc, and it is expected that they will be important in the exploitation of the higher frequencies that the war has made available.

The work "Very High-frequency Techniques" represents a summary of the methods, theories, and circuits used by the Radio Research Laboratory *that it is believed will be of general interest to radio engineers and physicists.* The material presented is based on knowledge gained from many sources. Much of it originated with the Radio Research Laboratory staff, but much of it also came directly or indirectly from the numerous NDRC, industrial, and service groups with whom the Radio Research Laboratory worked in close cooperation. Because of this cooperation, it has in general been found impracticable to attempt allocation of credit for individual developments, either among individuals or organizations. Attention is, moreover, directed to the fact that the Radio Research Laboratory engaged in tube development only to a minor extent; rather, it worked closely with tube-development groups in the laboratories of other contractors of Division 15 in evaluating their products. Hence, the portions of this volume that discuss tubes draw heavily on information originated by Bell Telephone Laboratories, Federal Telephone and Radio Corporation, General Electric Company, Litton Engineering Laboratories, Radiation Laboratory of M.I.T. (Division 14, NDRC), RCA Manufacturing Company, Westinghouse Electric and Manufacturing Company, and others that carried on tube development for Division 15 of NDRC. In addition, acknowledgement is due to various organizations that influenced the Radio Research Laboratory's technical thinking and hence contributed indirectly, though in an important way, to this book. The more important of these include the

Radiation Laboratory of M.I.T., the General Radio Company, the Naval Research Laboratory, the Special Projects Laboratory at ATSC, the Telecommunications Research Establishment (TRE) of England, and the many manufacturers of countermeasures equipment.

This volume has been compiled under the able editorship of Dr. Herbert J. Reich and the group of assistants listed on the title page, each of whom devoted appreciable time to editorial work. In addition valuable assistance of an editorial nature was supplied by George E. Hulstede, Robert Hammett, and Gunnar Hok. The authors of the different sections were in every instance closely associated with the work they are reporting, but the information that they have compiled should be regarded as the results of the efforts of teams, of which the authors were important, but not the only, members.

The material presented in "Very High-frequency Techniques" is a monument to the many engineers and scientists who contributed to the Radar Countermeasures program of Division 15 of the NDRC and in particular to the several hundred very able men and women whose technical experiences while on the staff of the Radio Research Laboratory are recorded in this volume.

F. E. TERMAN

Director, Radio Research Laboratory

CAMBRIDGE, MASS.,
April, 1946.

CONTENTS

Volume I

	PAGE
PREFACE	v
FOREWORD	vii
Chapter 1. BROAD-BAND ANTENNAS	1
<i>by Andrew Alford</i>	
Antennas and Their Band Widths—Parameters That Affect Band Width—Electromagnetic Horns—Antennas for Lower Frequencies: Impedance Compensation—Calculation of the Reflection Produced by Two Networks along a Transmission Line—Radiation Patterns and the Elementary Huygens' Principle—Equivalence Principle and Radiation Patterns of Apertures—Elementary Dipole; Vector-scalar-potential Method; Radiation Patterns of Thin Antennas—Radiation Patterns of Cylinders—Relation between Radiation Patterns of Conductors and Apertures—Some General Properties of Directional Antennas.	
Chapter 2. ULTRAHIGH-FREQUENCY MEASUREMENTS.	26
<i>by Jessie A. Nelson, David Lazarus, John W. Christensen, and Robert R. Buss</i>	
Transmission-line Measurements—The Smith Chart—Components of Impedance-measuring Equipment for Wide Frequency Ranges—Detector Probes—Use of Probe with Receiver—Calibration of Probes and System—Line-to-load Connections—Use of Cable between Measuring Line and Load—Measurement of Frequency—Use of Tuning Devices in Tuning the Generator—Elimination of Harmonics—Ground Plane—A Complete Measurement Setup—Measurement Procedure—Balanced-line Measurements Using Two Radio-frequency Cables—Waveguide Measurements: Slotted Sections—Probes, Detectors, and Indicators—Tuning Sections—Waveguide Loads—Measurement Procedure—Example of Waveguide Measurements—Antenna Pattern Measurements—Basic Pattern-measuring Techniques—Signal Sources—Detectors and Receivers—Test Antennas—Site and Mechanical Facilities—Automatic Pattern-measuring Methods.	
Chapter 3. IMPEDANCE MATCHING, TRANSFORMERS AND BALUNS.	53
<i>by Jessie Ann Nelson and Gus Stavis</i>	
Broad-band Impedance Compensation—The Definition Circle—Transformation Circles—Types of Compensating Networks—The Line Transformer: Preliminary Considerations, Transformation Circles—Impedance-curve Transformation by Line-transformer Methods—The Series Network—The Parallel Network—Compensation with the Parallel Network—General Considerations in Designing Single-element Broad-band Compensating	

Networks—Multielement Compensating Networks—Baluns: General, Type-I, Type-II, Type-III—Use of the Balun as Part of an Impedance-matching Network—Method of Checking Balun Performance.	
Chapter 4. CONE AND CYLINDER ANTENNAS	93
<i>by Arthur Dorne</i>	
General Considerations: Impedance, Patterns—Cone Antennas—Cylinders—Effect of Reflectors on Cones and Cylinders.	
Chapter 5. SLEEVE ANTENNAS	110
<i>by Ernest L. Bock, Jessie A. Nelson, and Arthur Dorne</i>	
General Considerations—Impedance and Pattern Characteristics of Cylindrical-sleeve Antennas—Examples of Sleeve Stubs—Simple Sleeve Dipoles—Bent-sleeve Dipoles—Sleeve Dipole-reflector Assemblies—Crossed-element Phased Antennas—Broad-band Crossed-element Phased Antennas	
Chapter 6. HORNS AND REFLECTORS	138
<i>by Gus Slavis and Arthur Dorne</i>	
Introduction—Methods of Feeding Horns—Radiation Patterns of Horns—Beam Width as a Function of Aperture—Circularly Polarized Horns—Measurements of Circularity—Reflectors: General Considerations—Elpar (Elliptical-paraboloid) Reflector.	
Chapter 7. SLOT ANTENNAS	171
<i>by Arthur Dorne and David Lazarus</i>	
Introduction—Radiation Patterns of Slot Antennas—Effect of Size of Ground Plane on Radiation Patterns—Effect of Width of Slot on Radiation Patterns—Impedance Characteristics of Slot Antennas—Methods of Reducing Size of Antenna—Examples of Slot Antennas.	
Chapter 8. AUXILIARY DEVICES FOR ANTENNAS	192
<i>by Ernst L. Bock and Joseph Margolin</i>	
Broad-band Microwave Switches: General Considerations—Coaxial Radio-frequency Switches—Flat-bladed Switches and Relays—Waveguide Switch.	

ture—Relation between Apertures—Power Gain—Directivity Ratio—Effect of Polarization on Aperture—Example of the Relation of the Apertures—Effect of Transmission Line on Effective Aperture—Band Width—Polarization Discrimination and Response Ratios—Bearing Deviation—Application of Direction-finder Antennas—Half-wave Element and Adcock Array—Double-loop Spinner: Combination Vertical-horizontal Corner-reflector Spinner—Combination Vertical-horizontal Parabolic-reflector Spinner—Spinner Auxiliaries—Circularly Polarized Horn with Parabolic Spinner—Fixed Antennas—Rotation-speed Considerations—Testing Techniques.

Chapter 11. INDICATORS FOR DIRECTION FINDERS. 272

by William McGuigan

The Problem—Remote Indication—Need for Automatic Direction Finding—Rotating-disk Indicator—Pulse Indicators—Magnetic-coil Method—Envelope Tracers—Magnetic Scan—Electrostatic Pulse Indicator—Capacitor Scan—Electrostatic Indicator—Sine-potentiometer Scan—Electrostatic Indicator—Electronic Scan—Radial Presentation—Electrostatic Indicator—Synchro Scan—Double-line Indicator—Comparison of Presentations—Reference Bearing—Instrument Errors.

Chapter 12. HOMING SYSTEMS 294

by John W. Christensen

Introduction to Homing Antennas for Very-high-frequency Homing—Pattern Measurement—Classification of Antenna Systems—Antenna Systems for Azimuth Homing—Antennas for Elevation Homing—Commutators—Indicators—Miscellaneous Considerations. A-v-c Circuits—Performance.

Chapter 13. INTRODUCTION TO POWER GENERATION. 311

by John F. Byrne

Military Requirements as to Frequency and Power—Tube Program—Modulation—Transmitting Equipment—Types and Requirements—Summary of Progress in the Field of Power Generation.

Chapter 14. TRIODE AND PENTODE ULTRAHIGH-FREQUENCY OSCILLATORS. . . 316

by Edwin A. Yunker, John F. Byrne, J. Gregg Stephenson, and Milton B. Adams

Introduction—Oscillator Analysis—Choice of Tubes and Circuits—Parallel-line Oscillators—450- to 720-Mc 8-watt Oscillator—250- to 550-Mc 20-watt Oscillator—Oscillator Using Lines of Variable Characteristic Impedance—Coiled Lines—Spiral-line Oscillator—Helical-line Oscillators—Notes on Design and Construction—Cathode Chokes and Tuned Lines.

Chapter 15. COAXIAL-LINE POWER AMPLIFIERS AND OSCILLATORS. 337

by William R. Rambo

Coaxial-line Oscillators and Amplifiers—Vacuum Tubes for Coaxial Circuits—Circuit Choices for Ultrahigh-frequency Use—The Common-grid Circuit—Multiple-mode Tuning—Mode Separation—Loaded-Q Considerations—Limitations on Tuning Range—Feedback Considerations—Special Properties of Coaxial-line Tank Circuits—Ultrahigh-frequency

Oscillator and Amplifier Operation—Grid-circuit Considerations—Circuit Tuning—Amplifier Operation—Practical Design Considerations—Tube and Circuit Selection—Choice of Line Characteristic Impedances—Physical Length of Lines—Circuit-tuning Mechanisms—Circuit-to-tube Connecting Systems—Feedback Design—Direct-current Circuit Design—Plate Feed and Hats—Output Coupling—Mechanical Design Notes—Trouble Shooting and Holes—Modification for Amplifier Use—Multitube Circuits.

Chapter 16. POWER OUTPUT COUPLING METHODS. 376

by J. Gregg Stephenson

Output Coupling to Antenna or Load—Effect of Standing-wave Ratio on the Output of Ultrahigh-frequency Transmitters—Untuned Coupling Loop—Loop Tuned with Series Stub—Balun or Balanced-unbalanced Transformer Type of Output Couplers: Inductively Coupled; Conductively Coupled; Capacitively Coupled—Capacitance-probe Coupling—Combined Probe and Loop Coupling—Direct Coupling—Series Output Coupler for Parallel-line Circuit.

Chapter 19. THE RESNATRON.	445
------------------------------------	-----

by William G. Dow and H. William Welch

The Resnatron—Structural Features of the Resnatron—Cathode Cavity—Anode Cavity—Electrode-structure Design—Anodes—Static Characteristics—Equivalent Circuit— Q of the Anode Cavity—Shunt-impedance Requirements—Parasitic Oscillations—Transit-time Effects—Feedback Circuit—Output Coupling—Transmission Line—Power Disposal—Operating Data—Modulation—Resnatron Transmitters.

Chapter 20. PRINCIPLES OF MAGNETRON OPERATION	473
---	-----

by Gunnar Hok and William G. Dow

Why and How Continuous-wave Magnetrons Were Developed—Relationship between Continuous-wave and Pulsed Magnetrons—Main Features of a Magnetron Oscillator—Classification of Continuous-wave Magnetrons—The Electromagnetic Field Configuration in a Multicavity Magnetron—Strapping—Circuit Constants of the Equivalent Network—The Electron Mechanism of Oscillation—Cathodes for Continuous-wave Magnetrons—Output Coupling—Achievement of Wide Tuning Range.

Chapter 21. OPERATING CHARACTERISTICS OF CONTINUOUS-WAVE MAGNETRONS.	502
--	-----

by Gunnar Hok

Introduction—Variation of Performance with Input Conditions: Performance Chart—Curves of Constant Magnetic Field: Mode-jump Current—Curves of Constant Frequency: Pushing—Series Magnet—Variations of Magnetron Performance with Load Impedance—The Rieke Diagram—Long-line Effect—Modulation.

Chapter 22. DESCRIPTION, CHARACTERISTICS, AND OPERATION OF TYPICAL CAVITY MAGNETRONS.	526
---	-----

by Gunnar Hok, H. William Welch, and William G. Dow

Introduction—Model 34 Magnetron—The 6J21 1-kw Tunable Magnetron—The Elimination of Extra Resonances—The Loaded Q as a Measure of Performance—Frequency Pushing—External Control of Band Width—Line Stretcher and Adjustable Reflector—Filament Emission and Back-heating—The Operating Characteristics of the 6J21—The 4J60-to-4J65 Low-power Magnetrons—The Squirrel-cage Magnetron or Donutron—Resonance Modes—Resonance Modes from Cold Tests— C -mode Operation—The Phase-reversing Anode—Mode Degeneracy—The Effect of Finger Shape on Tuning Range.

Volume II

Chapter 23. EXTERNALLY TUNED MAGNETRON OSCILLATORS	555
--	-----

by Herbert J. Reich

Externally Tuned Magnetron Oscillator Circuits—Types of Tube—Performance—Tube Life—Modulation—Load Coupling—Mode Jumping—Frequency Pulling—Frequency Pushing—Pulsing—Modulation by Filament Field.

	PAGE
Chapter 24. POWER-MEASURING DEVICES FOR ULTRAHIGH FREQUENCIES . . .	569
<i>by Edwin A. Yunker, Harold C. Early, Gunnar Hok, and Glendon R. Bridgeford</i>	
Introduction—Dummy Loads—Determination of Power from Current or Voltage and Resistance—Thermocouple in Coaxial Line—Disk Resistor with Voltmeter—Photometric Systems—Gas-filled Load Lamps—Calorimeter Wattmeters—Oil-immersed Dummy Load—Air Calorimeters—Lossy-coaxial-cable Calorimeter Wattmeter—Coaxial Line with Resistive Center Conductor—Coaxial Line with Water Dielectric—Salt-water Loads—Calorimeter Load for Waveguide—Quick-response Calorimeter for Large Waveguide—Sloping-guide Water-load Calorimeter—Slotted Coaxial Lines—Notched Coaxial Lines—The Thermistor—The Bolometer Bridge—Directional Couplers—Capacitive-loop Directional Coupler—Two-hole Directional Coupler—Bethe Hole Directional Coupler—A Broad-band Monitoring-type Wattmeter for Waveguides—Output Indicators—Spectrum Analyzers—Cold Analysis of Resonant Systems—Constants and Parameters of Resonant Systems—Equivalent Circuits—Measurements—Methods for Calculation of the Parameters—Analytical Foundations of Methods I to V.	
Chapter 25. RECEIVERS—GENERAL CONSIDERATIONS	627
<i>by Joseph M. Pettit</i>	
Design Objectives—Direct-detection Receivers—Superheterodyne Receivers—Definition and Measurements of Receiver Performance—Receiver Sensitivity—Noise Figure—Minimum Gain Requirement: Standard Noise Output—Signal Amplification: Standard Output—Combined Sensitivity Figure—Voltage Sensitivity vs. Power Sensitivity—Receiver Selectivity—Receiver Fidelity—Receiver Power Output Capability.	
Chapter 26. PRINCIPLES OF TRANSMISSION-LINE FILTER DESIGN . . .	648
<i>by Seymour Cohn</i>	
Fundamental Filter Relations—Calculation of the Image Impedance and Transfer Constant from the Short- and Open-circuit Impedances—Symmetrical-section Parameters in Terms of the Half-section Parameters—The Equivalent-circuit Method of Filter Analysis—Exact Equivalent-circuit Method of Filter Analysis—The Fundamental Analysis Method—Filter Insertion Loss—Filter Pass-band Response—Methods of Reducing Pass-band Insertion Loss—Transforming End Sections—Filter Stop-band Response—Pass-band Loss of a Filter with Improper Termination—Maximum Possible Insertion Loss of a Mismatched Filter—Line-loss Correction for Standing-wave Ratio—Waveguide as a Circuit Element in Filters—Properties of Ridge Waveguide—Higher Mode Cutoff Frequencies in Ridge Waveguide—Attenuation in Ridge Waveguide—Experimental Verification of the Ridge-waveguide Design Curves—Applications for Ridge Waveguide.	
Chapter 27. DESIGN OF TRANSMISSION-LINE FILTERS	685
<i>by Seymour Cohn</i>	
The Complete Filter—Short-line Filter Types—Short-line Low-pass Filter—Exact Design of the Varying-impedance Low-pass Filter—Design of a Typical Low-pass Filter—Short-line High-pass Filter— <i>M</i> -derived Termi-	

nating Half Sections for Low- and High-pass Filters—Short-line Band-pass Filters—Resonant-line Filter Types—Type 3-1, Band-pass and High-pass Filter—Type 3-2, Band-pass and High-pass Filter—Type 3-3, High-pass Filter—Type 3-4, Band-pass and High-pass Filter—Type 3-5, Low-pass Filter—Type 3-6, *M*-derived Band-pass Filter—The Filters of Tables 27-4 and 27-5—Mechanical Construction—Dielectric Materials—Effect of Physical Discontinuities—Specific Design—The Design of Simple Broad-band Waveguide-to-coaxial-line Junctions for Use with Waveguide Filters and Transmission Systems—The Basic Junction—The Transforming Junction—A Junction for Ridge Waveguide—Tapered-ridge Junction—Method of Testing the Waveguide-to-coaxial-line Junctions—Waveguide High-pass Filters—The Varying-impedance Band-pass Waveguide Filter—Test Methods for High-frequency Filters.

Chapter 28. TUNERS FOR MICROWAVE RECEIVERS. 741

by Seymour Cohn, Robert A. Soderman, Samuel J. Griffin, and Raymond O. Petrich

General Requirements for Receiver Tuners—The General Theory of Narrow-band Coupled-circuit Tuners—Losses in Radio-frequency Tuners—The Single-circuit Tuner—Two Coupled Circuits—Three Coupled Circuits—Butterfly Resonators: General Description—Butterfly Design—Butterflies as Tuners—Selectivity and Pass-band Loss—Coupling Methods—Spurious Responses—Examples of Butterfly Tuner Design—Double Butterfly—Mechanical Design—The Coaxial Cavity as a Tuner—Coupling Problems in a Coaxial Cavity—Coupling-loop Construction—Short-circuiting Devices—The Metal-to-metal Joint—Low-impedance Joint—The Choke Joint—Spurious Responses in Coaxial Cavities—Coupling between Coaxial Cavities—Design of Single-cavity Coaxial Tuners—Design of Two-cavity Coaxial Tuners—Design of Three-cavity Coaxial Tuners—Rectangular-waveguide Cavity—"Loaded" Waveguide Cavity—Rotary-tuned Coaxial Cavity—Rotary-tuned Waveguide Cavity—Movable-slug Coaxial Tuners.

Chapter 29. DETECTORS AND MIXERS 796

by Robert A. Soderman

Introduction—Types of Mixers and Detectors—Crystal Rectifiers—Crystal-mixer Circuits—Crystal-mixer Characteristics—Radio-frequency Impedance Considerations—Effect of Changes in the Local-oscillator Power Input on Mixer Performance—Local-oscillator Coupling Problems—Termination of Waveguide in Broad-band Untuned Waveguide Mixers—Crystal Characteristics Affecting Conversion Loss at High Frequencies—The Optimum Intermediate Frequency from Noise Considerations—Effect of Bias on Crystal-mixer Performance—Harmonic Mixing—Diode Mixers—Examples of Crystal and Diode Mixers—Low-level Direct Detection at High Frequencies—Diode and Crystal Low-level Detectors—Equivalent Circuit of a Low-level Crystal Detector—Noise in Low-level Crystal Detectors—Direct-detector Receiver Sensitivity—Input Impedance of a Low-level Detector—Examples of Wide-band Direct Detectors—Second Detectors—Crystals as Measuring Devices.

Chapter 30. LOCAL OSCILLATORS. I. GENERAL CONSIDERATIONS AND BUTTERFLY OSCILLATORS	824
<i>by Robert A. Soderman, William H. Huggins, and Frederick J. Kamphoefner</i>	
General Requirements—Survey of Oscillator Types—Butterflies as Resonators for Wide-range Oscillators—Oscillator Circuits—Butterfly Design—Grid-blocking Capacitor and Resistor Considerations—Plate-supply Connections—Sources of Holes—Effect of Cathode-lead Inductance—High-frequency Limitations—Stability—Output-coupling Methods—Butterfly Oscillator Efficiency—Oscillator Tubes—Examples of Butterfly Local Oscillators—Mechanical Tracking of Oscillator with Preselector and Dial.	
Chapter 31. LOCAL OSCILLATORS. II. REFLEX-KLYSTRON OSCILLATORS. . . .	849
<i>by William H. Huggins, Raymond O. Petrich, and Joseph W. Kearney</i>	
Introduction—The Electronic Admittance—Power Stability—Frequency Stability—The Resonator Admittance—Electronic Tuning—Oscilloscopic Techniques—Electronic Multiple Transits and Hysteresis—Harmonic-frequency Generation—Effect of Parasitic Resonances—Power-supply Requirements—Repeller-voltage Tracking Systems.	
Chapter 32. LOCAL OSCILLATORS. III. WIDE-BAND COAXIAL-LINE RESONATORS	878
<i>by William H. Huggins, Howard M. Zeidler, and Laurence A. Manning</i>	
Introductory Summary—Circular Coaxial Resonators—Summary of Basic Wave Types—Transverse Electromagnetic (<i>TEM</i>) Mode—Transverse Electric (<i>TE</i>) Modes—Transverse Magnetic (<i>TM</i>) Modes—Adaptation of Coaxial Resonators to Typical Reflex Tubes—Conditions for Oscillation—Equivalent Corner Length—Effect of Inductive and Capacitive Corner Terminations—Physical Realizability of λ - <i>D</i> Tuning Characteristic—Effect of Cavity Dimensions—Effect of Plunger Reactance—Repeller-voltage-vs-cavity-length Tuning Characteristics— <i>TEM</i> -mode Interference—Desirable Modes and Ranges—Elimination of <i>TEM</i> -mode Interference—General Analysis of <i>TEM</i> -mode Suppressor—Design of the $k = 1$ Mode-suppressor Section—Parasitic Circumferential Resonances of the Suppressor Section—Parasitic <i>TE</i> and <i>TM</i> Modes— <i>TE</i> Parasitic Resonances— <i>TM</i> Parasitic Resonances—Output-coupling Methods—Optimum Position of a Fixed Loop in $\frac{3}{4}$ -wavelength Resonator—Combination Loop-probe Methods—General Consideration of Noncontacting Plunger Design—Compound-line Section—Basic Noncontacting Plunger Types: Capacitance Plunger—Choke and Bucket Plungers— <i>S</i> and <i>Z</i> Plungers—Measurement of Plunger Reactance—Internal Power Loss of Plunger—Total Power Loss of Plunger—Permissible Plunger Loss in Reflex Oscillators—Parasitic Resonances in a Noncontacting Plunger.	
Chapter 33 INTERMEDIATE-FREQUENCY AMPLIFIERS	940
<i>by Matthew T. Lebenbaum</i>	
Introduction—Choice of Intermediate Frequency—Band Width and Selectivity Characteristics—Noise in Intermediate-frequency Amplifiers—Input Admittance of Vacuum Tubes—Gain—General Serviceability and Tube-constant Variations—Interstage Coupling Methods—Single-tuned Coupl-	

ing—Multituned Coupling Networks—Stagger Tuning—Design of a Stagger-tuned Amplifier—Regeneration—Gain Control—Automatic Gain Control.

Chapter 34. RECEIVER OUTPUT CIRCUITS 977

by Robert R. Buss, Jerre D. Noc, C. Bruce Clark, and Horace E. Overacker

Video Output—Band Width of Video Amplifiers—Control of Gain in Video Amplifiers—Measurement Problems—Video Output-coupling Circuits—Pulse Stretchers—Panoramic Presentation—Spurious Responses—Marker Signals for Frequency Indication—Phase and Amplitude Response of the Video and Deflection Amplifiers—Pulse Stretching and Trace Intensification—Recording Circuits and Equipment—Pulse Analysis—Oscilloscopic Measurement of Pulse Length—Meter Indication of Pulse Length—Oscilloscopic Measurement of Pulse-repetition Frequency—Meter Indication of Pulse-repetition Frequency.

Chapter 35. MEASURING EQUIPMENT FOR RECEIVERS 1008

by W. Bruce Wholey

Test Equipment—Signal Sources—Components of Signal Sources—Radio-frequency Oscillators—Modulators—Pulse Modulation of Radio-frequency Oscillators: Lighthouse Tubes, Reflex Klystrons—Output Systems—Piston Attenuators—Attenuation Pads—Output Calibration—Power Measurement—Shielding—Auxiliary Output Systems—Ultrahigh-frequency Sweep Oscillator—Heterodyne Frequency Meters.

Bibliography. SUPPLEMENTARY BIBLIOGRAPHY OF RECENT ARTICLES IN THE FIELD 1033

INDEX 1037

CHAPTER I

BROAD-BAND ANTENNAS

By ANDREW ALFORD

1-1. Antennas and Their Band Widths.—An antenna is a transformer for converting guided electromagnetic waves into free waves. In general, the relations between the electric and magnetic vectors and linear dimensions in a guided wave are different from those in a wave in free space, and therefore it is necessary to provide some means of impedance matching between waves in free space and waves in a transmission line or waveguide. The frequency range over which the transformation between guided waves and free-space waves can be effected depends on the nature of the mismatch in this transformer. If the transformation has the nature of a gradually tapered line so that the controlling parameters do not undergo a sudden change but vary gradually between a point where the waves are definitely guided waves and a point where the waves are definitely free-space waves, then, as in other gradually tapered transitions, the reflections are distributed over a distance of one or several wavelengths and have a chance partly to cancel because of phase differences. In a tapered transition the over-all reflection is not critically dependent either on the exact length of the transition section or on the exact wavelength of the transmitted waves.

When the transformation between guided and free-space waves is in the nature of a sudden change, there is a region of reflection such that a portion of the power is sent back into the line from which it came. This reflection can be counteracted by an equal reflection somewhere in the line that produces a reflected wave equal in amplitude but opposite in phase to that which is transmitted back through the second reflection point from the first reflection point. In such an arrangement the distance between the two reflection points plays a major part. When this distance is substantial in terms of the wavelength, a slight change in frequency changes the relative phase of the two reflected waves so that they no longer cancel out. A transformer in which the two reflections occur becomes a frequency-sensitive circuit with a frequency region within which the reflections cancel, but outside of which it no longer acts as a matched transformer. Under these circumstances the transformer is said to have a band width. Within the band width the transformer produces a degree of match that can be described by the amplitude of

the reflected wave when the direct wave is assumed to have unit amplitude. The band width is a relative term that can be made definite only by specifying the maximum value of the reflection coefficient not to be exceeded within the band. Thus, if we specify that the amplitude of the reflected wave shall never exceed one-third of the amplitude of the direct wave, the band width has a definite value, characteristic of different transformers. These band widths on the basis of a reflection coefficient of $\frac{1}{3}$ are, of course, greater than the band widths corresponding to a reflection coefficient of, say, $\frac{1}{10}$. Since various applications of antennas call for different maximum values of the reflection coefficient, the definitions of band width vary.

One commonly used definition of band width is based on a reflection coefficient equal to one-third. Another frequently used definition calls for the maximum value of reflection coefficient to be less than $\frac{1}{10}$. The first is useful when antennas are employed with transmitters for communication and other purposes, while the second is applied to radar and directional communication antennas consisting of a large number of individual elements. Still another definition is commonly used in connection with receiving antennas. This definition calls for a reflection coefficient ρ less than $\frac{2}{3}$.

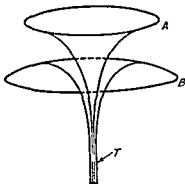


FIG. 1-1.—A broad-band antenna.

The term *broad-band antenna*, like the band width, is also relative. The general concept is that an antenna having a large band width is a broad-band antenna. A half-wave antenna made of thin wire for lower ultrahigh frequencies has a frequency band width of the order of 15 per cent on the basis of $\rho = \frac{1}{3}$. This antenna is regarded as a narrow-band antenna. The term broad-band antennas is therefore reserved for those which have band widths of over 15 per cent, on the same basis. This is an arbitrary definition. It is obvious that it can be challenged, for it amounts to making exact by definition a concept that is by nature quite vague. It is probably safer to say that a broad-band antenna is one that has a band width substantially greater than that of a thin-wire half-wave antenna, whatever the basis of comparison may be.

Perhaps the simplest broad-band antenna is that shown in Fig. 1-1. In this figure a concentric transmission line T gradually diverges in such a way that the ratio between the diameters of the inner and outer conductors is kept constant. The law of divergence is roughly exponential. If it is assumed that the change in dimensions per wavelength in the tapered transmission line is not very great, there will be relatively little

reflection at any particular point along the diverging line. A wave traveling along the line will make the gradual bend, will expand over larger and larger space, and finally having reached the end of the line will proceed into free space. If the distance between the edges *A* and *B* in Fig. 1-1 is of the order of a wavelength or more, very little reflection occurs in the region of departure, *i.e.*, along the surface of revolution that separates free space from the region between the two expanded conductors of the transmission line.

The structure in Fig. 1-1 is an antenna, though with equal justification it can be called a *horn*. Experiment shows that there is nothing

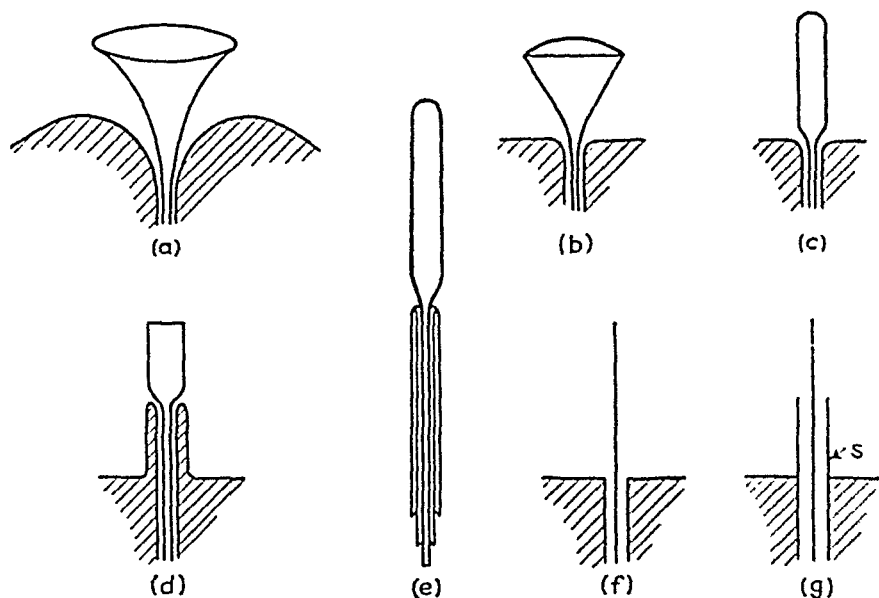


FIG. 1-2.—Antennas that may be regarded as distortions of the antenna shown in Fig. 1-1.

very critical about such a structure. It may be distorted to a considerable extent without losing its main property of being a wide-band transformer between a concentric transmission line and free space if, in the process of distorting the structure, the feature of small change in dimensions per wavelength is kept. If the structure is distorted so that there is a sudden change in cross section or in any other feature, one or more sources of reflection are introduced and the band width is reduced. Figure 1-2 illustrates a number of ways in which the structure of Fig. 1-1 may be distorted.

1-2. Parameters That Affect Band Width.—In order to gain an idea of some of the parameters that affect the band width of the various distorted antennas of Fig. 1-2, it is convenient to consider a family of cone antennas, as shown in Fig. 1-3. One characteristic that distinguishes a cone antenna from some of the other antennas is the fact that the cone

together with the flat plate constitutes a transmission line having a definite value Z_c of characteristic impedance. This value is

$$Z_c = 138 \log_{10} \cot \frac{\Psi}{2} \quad (1-1)$$

Figure 1-4 shows the value of Z_c as a function of the angle of flare Ψ .

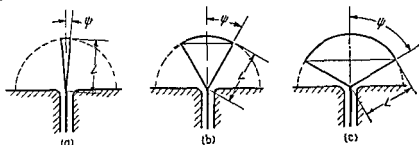


FIG. 1-3—Cone antennas with different values of characteristic impedance Z_c .

In general, the characteristic impedance Z_0 of the transmission line which is used to carry power to or from a cone antenna is different from the characteristic impedance of the cone, so that there are two reflection zones, one at the aperture, i.e., in the region near the surface S indicated

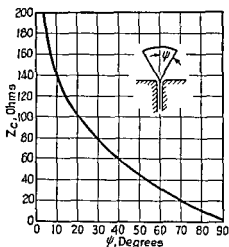


FIG. 1-4.—The relation between the characteristic impedance and the angle of flare of a cone antenna

by dotted line in Fig. 1-3, and another at the apex of the cone. In this respect a cone antenna is typical of all the antennas of Fig. 1-2 except antenna a . Other antennas, such as cylinders or cone cylinders, all possess this double-reflection feature. In cylinders the phenomena are complicated by the fact that the characteristic impedance is not constant but increases with the distance from the point of junction with the transmission line, so that there is an additional region of distributed reflections.

If the length L of the cone is made slightly less than $\frac{1}{4}$ wavelength and the cone is very thin,

angle Ψ being of the order of 1 deg, the antenna is, in effect, a thin-wire quarter-wave antenna. The base impedance of such an antenna is a resistance of the order of 36 ohms. The value of base impedance is thus much lower than the 284-ohm (for $\Psi = 1$ deg) characteristic impedance Z_c of the cone. This is true because of the relatively high

value of the reflection at the aperture, *i.e.*, in the region of the end of the antenna. Let the current reflection coefficient at the aperture be $\rho = \rho_0 e^{i\alpha}$. Then at the current maximum, approximately $\lambda/4$ from the end, *i.e.*, at the apex of the cone, the current is $I_0(1 + \rho_0)$, while the potential is $I_0 Z_c(1 - \rho_0)$. The input impedance at the apex is

$$Z_c \frac{1 - \rho_0}{1 + \rho_0} = 36 \quad \text{ohms}$$

Hence

$$\rho_0 = \frac{1 - \frac{36}{Z_c}}{1 + \frac{36}{Z_c}} = 0.775$$

If the cone angle is increased, the characteristic impedance Z_c rapidly decreases, as shown in Fig. 1-4. The value of ρ_0 also decreases and at such a rate that $Z_c(1 - \rho_0)/(1 + \rho_0)$ remains roughly constant. As the characteristic impedance Z_c of the conical transmission line approaches a value similar to the characteristic impedance of convenient concentric transmission lines, it just so happens that at the same time the aperture reflection coefficient ρ_0 approaches a low value. When the cone characteristic impedance is decreased further, the aperture reduces to a narrow ringlike slot, which presents a large mismatch with space, and ρ_0 increases again. In the intermediate region the cone transmission line has a convenient value of Z_c and is terminated "at the aperture" in an impedance nearly equal to Z_c .

Since the dimensions of the aperture in wavelengths are changed when frequency is varied, the aperture reflection coefficient of a given cone varies rapidly with frequency. Typical behavior of ρ_0 is shown in Fig. 1-5. If the input impedance of a cone antenna is plotted on a resistance-reactance diagram with one point for each frequency, the resulting curve is a spiral that curls around the value of Z_c (see Fig. 1-6). This spiral shape is explained as follows. The impedance at a distance θ deg from the end of a transmission line terminated by an impedance Z having

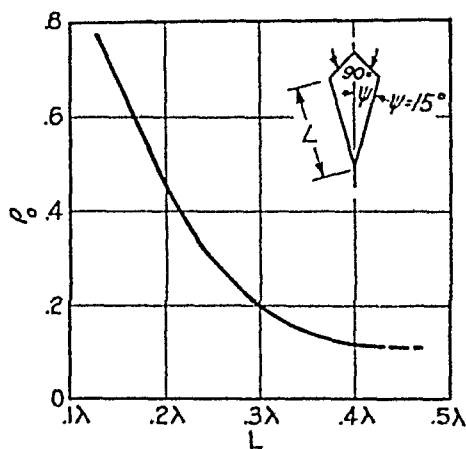


FIG. 1-5.—Approximate value of the coefficient of "aperture" reflection. ρ_0 is the ratio of the amplitude of the reflected wave to the amplitude of the outgoing wave as measured at the apex of the cone.

coefficient of reflection $\rho = \rho_0 e^{j\alpha}$ is $Z_1 = R + jX$ where

$$R = \frac{1 - \rho_0^2}{1 + \rho_0^2 + 2\rho_0 \cos u} Z_e, \quad X = + \frac{2\rho_0 \sin u}{1 + \rho_0^2 + 2\rho_0 \cos u} Z_e \quad (1-2)$$

and

$$u = 2\theta - \alpha$$

If u is eliminated from these equations, the result is

$$\left(R - Z_e \frac{1 + \rho_0^2}{1 - \rho_0^2}\right)^2 + X^2 = \left(Z_e \frac{2\rho_0}{1 - \rho_0^2}\right)^2 \quad (1-3)$$

This is a circle of radius $Z_e \frac{2\rho_0}{1 - \rho_0^2}$ with the center located at

$$R = Z_e \frac{1 + \rho_0^2}{1 - \rho_0^2}, \quad X = 0.$$

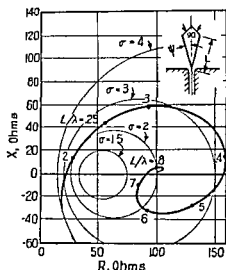


FIG. 1-6—A typical R - X diagram of the impedance of a cone antenna. The numbers associated with dots along the impedance curve indicate the values of L in wavelengths that correspond to the values of X and R represented by the points. The numbers associated with circles indicate the values of the standing-wave ratio that would result if the transmission line were terminated by any impedance represented by a point on the circle.

This result, when applied to a cone antenna, leads to the following conclusions. At lower frequencies the value of ρ_0 is large, so that the impedance point (R, X) lies on a circle of a large radius with the center at a large distance from the axis $R = 0$. As the frequency is increased, the radius of the circle $Z_e \frac{2\rho_0}{1 - \rho_0^2}$ decreases while the center of the circle moves closer to $R = Z_e$. Thus with decreasing ρ_0 the impedance point moves along a curve with a smaller and smaller radius of curvature. When ρ_0 is very small, the radius of the circle is also very small and the center of the circle is near Z_e . The over-all result is a spiral, as shown in Fig. 1-6.

So far we have outlined the behavior of a cone radiator. Other radiators of shapes intermediate between the radiator of Fig. 1-1 and the radiator of Fig. 1-2f behave in a similar manner. Fat cylinders are no exceptions in this respect. If a cylinder is provided with a conical taper of proper Z_e near the plate that serves as the return conductor, the localized excess of capacitance is reduced to the point where there is only a slight

tendency for the impedance spiral to be shifted in the direction of negative reactance because of the lumped base capacitance. When a taper is not used, the tendency of the spiral to be displaced in the negative-reactance direction is very marked. Except for this detail, the R - X impedance diagram is a spiral, as in Fig. 1-7, that is very similar in shape to that of the cone shown in Fig. 1-6. Thin structures, such as the one in Fig. 1-2g, do not exhibit a rapidly decreasing aperture reflection coefficient as the frequency is increased. On the contrary, the reflected wave proceeding in the direction away from the end of such an antenna remains equal to a substantial fraction of the direct outward wave. The low

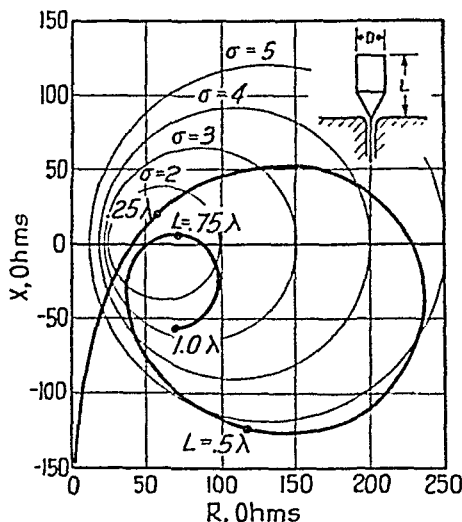


FIG. 1-7.—A typical R - X diagram of a cylindrical antenna. $D = 1/10 L$.

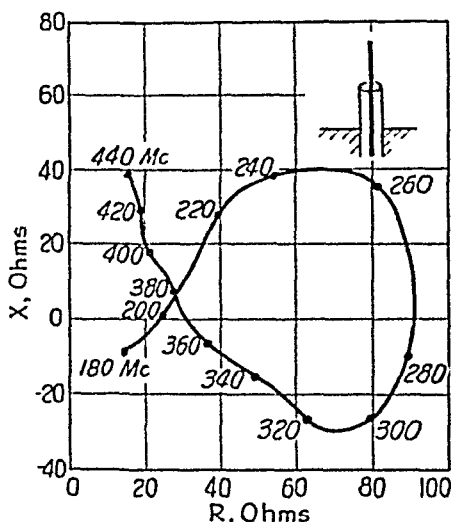


FIG. 1-8.—A typical R - X diagram of a sleeve antenna.

value of the reflection coefficient at the input is obtained by providing one or more compensating reflections within the antenna itself or in an associated compensatory circuit.

In the case of the antenna in Fig. 1-2g, the major portion of the compensation is provided by the reflected wave traveling upward along the sleeve s . This wave results in an inductive impedance between the end of the outer conductor and ground, while the reflection at the end of the inner conductor and ground is capacitive. Since the two impedances are effectively in series with each other, the reactive components tend to cancel, leaving a resistive termination for the feeder.

As in other thin antennas, a large reflection may be compensated for by another, correspondingly large, reflection, if the two reflections are in proper relative phase to reduce the total reflection to a low value. This compensation occurs only within a limited band of frequencies. Outside this band the two reflections compound to produce a total reflection that

is greater than either of the individual reflections. The over-all result is that thin antennas with substantial band widths behave somewhat like band-pass filters with steep cutoff characteristics at the lower and upper limits of the pass band. The corresponding R - X curve has a large radius of curvature at lower frequencies, then a small radius of curvature corresponding to the lower values of ρ_0 within the pass band, and finally a large radius of curvature above the pass band. The resulting shape of the R - X curve is a loop, as shown in Fig 1-8. This curve is typical of a structure in which the aperture reflection is relatively large but is compensated for by another reflection within a limited band of frequencies.

1-3. Electromagnetic Horns.—A gradual transition between a waveguide and space can be accomplished in much the same way as the transition between a concentric transmission line and space. Figure 1-9 shows the so-called *electromagnetic horn*, which is a wide-band transformer

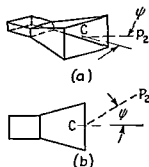


FIG. 1-9.—A horn antenna.

between guided waves and waves in space. Here, as in Fig. 1-1, the parameters affecting the propagation vary gradually at a slow rate so that there is only a small change per wavelength from a point that is in the waveguide to a point that is in free space. If the aperture of the horn is wide enough, a wavelength or more, and the angle of flare, *i.e.*, the rate of taper, is not too great, the transition between guided waves and free space is effected with very little reflection. Again, as in the antenna of Fig. 1-1, the exact shape of the flare or the exact dimensions

of the aperture are not critical. The rate of flare, *i.e.*, the change in cross-sectional dimensions per wavelength, may also be varied over wide limits without much effect, provided it is sufficiently small to start with. The shape of the horn may be distorted in various ways, but as long as the feature of small change in cross-sectional dimensions per wavelength is preserved, the band width of the antenna remains great.

From the point of view of impedance, the electromagnetic horn is in some respects very similar to a cylindrical antenna, shown in Fig. 1-2c. At the aperture there is a zone of reflection. The horn itself is a waveguide with variable characteristic impedance. At the junction of the horn and the waveguide there is another discontinuity in characteristic impedance. These three zones of reflection are reminiscent, respectively, of the end reflection, of the variable characteristic impedance, and of the region in the immediate vicinity of the junction with the transmission line in the case of a cylindrical antenna.

When electromagnetic horns are used at lower frequencies, they are usually distorted to a considerable degree from the form shown in Fig.

1-9. One form of distorted horn is obtained by increasing the flare angle to 180 by 180 deg. This is particularly convenient when there is already a metal sheet that can serve as such a horn. A simple example of a distorted horn of this type is shown in Fig. 1-10. The principal source of reflection in this case is in the region where the waveguide joins with the horn, *i.e.*, with the metal sheet. The reflection coefficient in this case depends on the dimensions of the waveguide. With an ordinary rectangular $TE_{1,0}$ -mode waveguide with $b/a = 2$, the reflection coefficient is of the order of $\frac{1}{4}$. With a

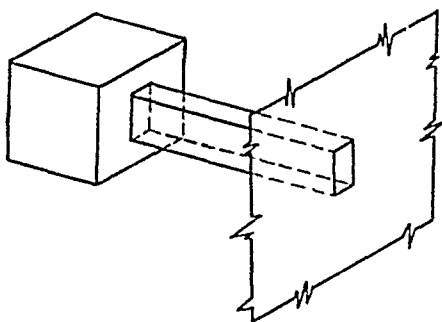


FIG. 1-10.—A rectangular opening in a large metal sheet. The opening is energized by a waveguide.

narrow waveguide having $b/a = 8$, the reflection coefficient is about $\frac{2}{3}$.

Because of the large size of waveguide for lower frequencies, it is more convenient to use coaxial transmission line for conveying power of such frequency to or from an antenna and to install a coaxial-line-to-waveguide transformer near the antenna. An arrangement of this type is illustrated in Fig. 1-11. Since a reflectionless transformation between coaxial line and waveguide would allow a wave reflected at the aperture to continue as a reflected wave into the transmission line, some effort is usually made to arrange for a compensating reflection in the coaxial-line-to-waveguide transformer. When this is done, the over-all structure acquires the characteristics of a band-pass filter. The shorter the distance between the aperture reflection and the compensating transformer reflection, the greater is the attainable band width. Also, the smaller the aperture reflection, the greater is the band width. Since the aperture reflection increases

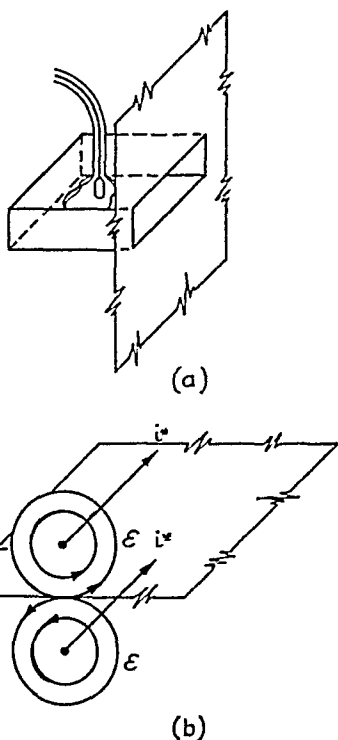


FIG. 1-11.—(a) A slot antenna. (b) The electric lines of force around two parallel magnetic currents flowing in the same direction showing the cancellation of the electric field along the plane equidistant from the currents.

when the width of the waveguide is decreased, the band width is smaller with narrow apertures.

The 180 by 180-deg horns provided with a short length of waveguide and with the coaxial-line-to-waveguide transformer mounted close to the aperture are often referred to as *slot antennas*. The impedance characteristics of slot antennas are quite similar in major respects to those of other compensated antennas. Below the frequency at which compensation is

effective, the reflection coefficient ρ_0 is large, so that the R - X curve has a large radius of curvature. Within the compensated band, ρ_0 is small, resulting in one or more small loops. At frequencies above the compensated band, ρ_0 is again large and the R - X curve proceeds with a large radius of curvature. When the compensation within the pass band is good, the value of ρ_0 is comparable to that of small reflections from the edges of the sheet, the reflections from connectors and from other sources. These, collectively, sometimes contribute small curls in the R - X curve. All

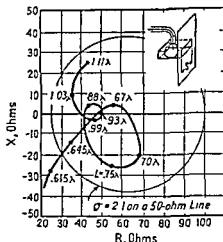


FIG. 1-12.—A typical R - X diagram of a broad-band slot antenna.

these effects are illustrated in Fig. 1-12, which shows a typical R - X curve of a slot antenna.

1-4. Antennas for Lower Frequencies: Impedance Compensation.—

In practice it is almost always desirable to keep the volume occupied by any antenna to the very minimum. Structures like that shown in Fig. 1-1 are, as a rule, uneconomical in the sense that they occupy more space than other structures that might have the required band width. Smaller structures are obtained by allowing a greater reflection in the transformation between guided waves and space waves, but compensating it by a reflection in the transmission line. The lower the frequency, the more important it is to make use of various methods of compensation in order to avoid the great bulk that a broad-band radiator is otherwise likely to have. At the lower frequencies the gradual transitions illustrated in Figs. 1-1 and 1-9 result in very unwieldy dimensions and are therefore generally avoided. The design of broad-band antennas for

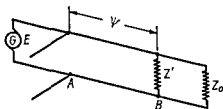


FIG. 1-13.—A transmission line terminated by its characteristic impedance Z_0 and with two reflection points, one at A , the other at B , each reflection produced by a shunt impedance.

frequencies below 700 Mc or so almost always calls for some form of compensation. It is appropriate, therefore, that we discuss in broad terms a typical example of compensation.

To make the discussion more concrete, it is convenient to consider the arrangement shown in Fig. 1-13, in which a transmission line is terminated in its characteristic impedance Z_0 and is provided with two reflection points, one at A and another at B , which are spaced along the transmission line by a distance of Ψ deg. This arrangement corresponds to an antenna provided with a compensating device which is placed Ψ deg from the antenna. The characteristic impedance Z_0 in Fig. 1-13 corresponds to the free-wave propagation into space; the first reflection at B corresponds to the reflection produced in the antenna; the second reflection at A is the reflection introduced by the compensating circuit. This is an elementary arrangement that will be used to illustrate two important points: (1) The greater the distance of the compensating circuit from the antenna, the narrower is the band width. (2) The greater the individual reflections at A and B , the narrower is the band width.

In order to prove these points, it is convenient to introduce several auxiliary concepts. Let a transmission line terminated by its characteristic impedance be shunted by a reactance jX . Let the current amplitude of the direct wave that travels from the generator toward the reactance be unity, and let the phase of this wave as it arrives at the junction point where the reactance is shunted be zero. Let ρ , a complex number, represent the amplitude and phase of the reflected wave as it leaves the junction point. Let τ , also a complex number, represent the amplitude and phase of the transmitted wave as it leaves the junction point. The following notation is convenient:

$$\begin{aligned}\rho &= \rho_0 e^{j\alpha} \\ \tau &= \tau_0 e^{j\eta}\end{aligned}$$

Along the transmission line between the generator and the shunted impedance there are standing waves. At a point where current maximum occurs the current is $(1 + \rho_0)$ and the potential is $Z_0(1 - \rho_0)$. Since at such a point the current and voltage are in phase, the transmitted power is

$$P = Z_0(1 - \rho_0)(1 + \rho_0) = Z_0 - Z_0\rho_0^2 \quad (1-4)$$

The power that is delivered to the load at the end of the transmission line is $\tau_0^2 Z_0$. Since there is no loss of power in the reactance, we have

$$Z_0 - Z_0\rho_0^2 = \tau_0^2 Z_0$$

Hence

$$\tau_0^2 + \rho_0^2 = 1 \quad (1-5)$$

On the transmitter side of the shunt the potential is $Z_0(1 - \rho)$. Since there is no generator at the shunt, potentials on both sides of the shunt are equal, so that

$$Z_0(1 - \rho) = Z_0\tau$$

Therefore

$$\tau + \rho = 1 \quad (1-6)$$

This is a vector equation in which both numbers are complex.

At the junction point the impedance is Z , given by

$$\frac{1}{Z} = \frac{1}{jX} + \frac{1}{Z_0} \quad (1-7)$$

The current-reflection coefficient

$$\rho = \frac{Z_0 - Z}{Z_0 + Z} = \frac{\frac{Z_0}{jX}}{\frac{Z_0}{jX} + 2} \quad (1-8)$$

In the special case for which the shunting reactance is a length of Δ deg of open-circuited transmission line of characteristic impedance Z_0 ,

$$X = -Z_0 \cot \Delta \quad \text{and} \quad \rho = \frac{-\tan \Delta}{-\tan \Delta + 2j} \quad (1-9)$$

Let θ be an auxiliary length in degrees defined by the equation

$$\tan \Delta = 2 \tan \theta \quad (1-10)$$

Then

$$\rho = \frac{-\tan \theta}{-\tan \theta + j} = \sin \theta e^{j(\frac{\pi}{2} - \theta)} \quad (1-11)$$

and

$$\tau = 1 - \rho = \cos \theta e^{-j\theta} \quad (1-12)$$

By comparing these results with the definitions of ρ and τ , we see that in this case

$$\left. \begin{aligned} \rho_0 &= \sin \theta, & \alpha &= \frac{\pi}{2} - \theta \\ \tau_0 &= \cos \theta, & \eta &= -\theta \end{aligned} \right\} \quad (1-13)$$

1-5. Calculation of the Reflection Produced by Two Networks along a Transmission Line.—Consider now the case of two points of reflection, as shown in Fig. 1-13. Let the distance between these points be Ψ deg. It will be convenient to work in terms of τ_1 and ρ_1 which are, respectively, the transmission and reflection coefficients of reactance at A , under the assumption that the reactance at A is the only one shunted across the transmission line. Similarly, let ρ_2 and τ_2 be, respectively, the

reflection and transmission coefficients of the reactance at B , also under the assumption that the reactance at B is the only one shunted across the transmission line. Let the amplitude of the current in the incident direct wave, which travels from the generator toward A , be unity, and let the phase of this current be zero at A . Let the *total* current in the forward wave as it leaves A in the direction of B be represented by the complex number U . As this direct wave arrives at B , it is equal to $Ue^{-j\Psi}$. The total reflected wave that travels from B to A is therefore $\rho_2 Ue^{-j\Psi}$. This wave is again reflected at A and forms a part of the direct wave U . The wave reflected at A toward B is $\rho_1 \rho_2 Ue^{-2j\Psi}$. The direct wave U is a sum of the portion τ_1 of the primary wave transmitted through the junction and the reflected wave $\rho_1 \rho_2 Ue^{-2j\Psi}$. Thus we have

$$U = \tau_1 + \rho_1 \rho_2 Ue^{-2j\Psi} \quad (1-14)$$

This equation can be solved for U and

$$U = \frac{\tau_1}{1 - \rho_1 \rho_2 e^{-2j\Psi}} \quad (1-15)$$

The total reflected wave R which goes back toward the generator is a sum of two waves, the wave reflected at A , which is equal to ρ_1 , and the portion of the reflected wave from B , which is transmitted through the junction at A , *i.e.*, $\tau_1 \rho_2 Ue^{-2j\Psi}$. Therefore,

$$R = \rho_1 + \frac{\tau_1^2 \rho_2 e^{-2j\Psi}}{1 - \rho_1 \rho_2 e^{-2j\Psi}} \quad (1-16)$$

By substituting

$$\begin{aligned} \rho_1 &= \rho_{10} e^{j\alpha_1}, & \rho_2 &= \rho_{20} e^{j\alpha_2} \\ \tau_1 &= \tau_{10} e^{j\eta_1} \end{aligned}$$

and using the fact that

$$\tau_{10}^2 + \rho_{10}^2 = 1 \quad \text{and} \quad \eta_1 - \alpha_1 = -\frac{\pi}{2}$$

we get

$$R = e^{j\alpha_1} \frac{\rho_{10} - \rho_{20} e^{j(\alpha_1 + \alpha_2 - 2\Psi)}}{1 - \rho_{10} \rho_{20} e^{j(\alpha_1 + \alpha_2 - 2\Psi)}} \quad (1-17)$$

The square of the absolute value of R is therefore

$$|R|^2 = \frac{\rho_{10}^2 + \rho_{20}^2 - 2\rho_{10}\rho_{20} \cos v}{1 + \rho_{10}^2 \rho_{20}^2 - 2\rho_{10}\rho_{20} \cos v} = \frac{\frac{(\rho_{10} - \rho_{20})^2}{(1 - \rho_{10}\rho_{20})^2} + \frac{4\rho_{10}\rho_{20}}{(1 - \rho_{10}\rho_{20})^2} \sin^2 \frac{v}{2}}{1 + \frac{4\rho_{10}\rho_{20}}{(1 - \rho_{10}\rho_{20})^2} \sin^2 \frac{v}{2}} \quad (1-18)$$

where

$$v = \alpha_1 + \alpha_2 - 2\Psi$$

The value of $|R|$ is minimum or maximum when $v = 0$ or $v = \pi$.
The maximum value can be found from

$$|R|^2 = \frac{(\rho_{10} + \rho_{20})^2}{(1 + \rho_{10}\rho_{20})^2} \quad (1-19)$$

the minimum, from

$$|R|^2 = \frac{(\rho_{10} - \rho_{20})^2}{(1 - \rho_{10}\rho_{20})^2} \quad (1-20)$$

The values of the standing-wave ratio that correspond to these values of $|R|$ are

$$\left. \begin{aligned} \sigma_{\max} &= \frac{1 + \left| \frac{\rho_{10} + \rho_{20}}{1 + \rho_{10}\rho_{20}} \right|}{1 - \left| \frac{\rho_{10} + \rho_{20}}{1 + \rho_{10}\rho_{20}} \right|} = \sigma_1 \sigma_2 \\ \sigma_{\min} &= \frac{1 + \left| \frac{\rho_{10} - \rho_{20}}{1 - \rho_{10}\rho_{20}} \right|}{1 - \left| \frac{\rho_{10} - \rho_{20}}{1 - \rho_{10}\rho_{20}} \right|} = \begin{cases} \frac{\sigma_1}{\sigma_2} & \text{if } \rho_{10} > \rho_{20} \\ \frac{\sigma_2}{\sigma_1} & \text{if } \rho_{10} < \rho_{20} \end{cases} \end{aligned} \right\} \quad (1-21)$$

where $\sigma_1 = \frac{1 + \rho_{10}}{1 - \rho_{10}}$, $\sigma_2 = \frac{1 + \rho_{20}}{1 - \rho_{20}}$. Thus we see that the standing-wave ratio that is produced by two networks must fluctuate between the product $\sigma_1 \sigma_2$ and the ratio σ_1/σ_2 or σ_2/σ_1 of the standing-wave ratios produced by these networks when each is placed alone on a reflectionless transmission line. Complete compensation, i.e., $\sigma = 1$, can be had only when $\sigma_1 = \sigma_2$ and $v = 0, 2\pi$, etc.

Let us assume that $\sin v/2 = 0$ and $\rho_{10} = \rho_{20}$ at the frequency $f = f_0$. Then at other frequencies near f_0 smaller values of $|R|^2$ are obtained if ρ_{10} and ρ_{20} remain equal, because the numerator of Eq. (1-18) is a sum of two positive terms. When ρ_{10} remains equal to ρ_{20} in the neighborhood of f_0 , where $\sin v/2$ is small, we have

$$|R|^2 \approx \left(\sigma_1 - \frac{1}{\sigma_1} \right)^2 \sin^2 \frac{v}{2}$$

and

$$\sigma \approx 1 + 2 \left(\sigma_1 - \frac{1}{\sigma_1} \right) \left| \sin \frac{v}{2} \right| = 1 + 2 \left(\sigma_1 - \frac{1}{\sigma_1} \right) \left| \sin \left(\frac{\alpha_1 + \alpha_2}{2} - \Psi \right) \right| \quad (1-22)$$

The rate at which σ deviates from unity thus depends on the product of $\left(\sigma_1 - \frac{1}{\sigma_1} \right)$ and $\sin \left(\frac{\alpha_1 + \alpha_2}{2} - \Psi \right)$. When Ψ , the distance between the networks, is large in comparison with the length of the compensating networks and the latter are not sharply resonant, as is usually the case,

the value of Ψ varies much more rapidly than $\alpha_1 + \alpha_2$. Then the deviation from unity of σ , the standing-wave ratio due to the two networks, is rapid. This rate of deviation is still further increased when σ_1 is large.

In deriving these equations, it has been assumed that the impedance shunted at A is a reactance. Nothing has been assumed in regard to the nature of the impedance at B except that it has a reflection coefficient $\rho_{20}e^{i\theta_2}$, which imposes no limitations at all. It can be shown that all the results derived above apply equally well when the compensating network on the generator side is a series reactance or, in fact, any nondissipative network or a combination of nondissipative networks. The proof of this more general theorem is lengthy and will not be given here.

The above calculations show that (1) at any one frequency it is always possible to cancel completely one reflection by another; (2) as the frequency is changed from this critical frequency, the reflection increases at a rate that is a function of the distance between the two sources of reflection; (3) the greater the distance between the two reflections, the greater is the rate of increase in over-all reflection; (4) the greater the individual reflections, the greater is the rate at which the total reflection increases with the deviation from the critical frequency. These are prevalent ideas. They apply to almost every kind of compensation that is employed in the design of broad-band antennas.

We have considered a simple case consisting of a single compensating reflection. A number of other possibilities exist. It is convenient to regard the pair consisting of the primary reflection and the first compensating reflection as one single source of reflection. This reflection produces zero reflected waves at the critical frequency but above and below this frequency results in a back wave having a certain amplitude and a certain phase. It is possible to insert into the transmission line one or a combination of several reflections that are zero at the critical frequency but that tend to cancel out the reflections of the first pair above and below the critical frequency. This might be called *compound compensation*. Compound compensation has been applied at frequencies below 70 Mc where other means for obtaining the desired band width resulted in prohibitive size of the antenna.

1-6. Radiation Patterns and the Elementary Huygens' Principle.—So far we have considered in some detail the phenomena that take place on the transmission-line, or waveguide, side of an antenna. We shall now inquire into what happens to the waves in free space after they leave the antenna. Consider first the horn radiator shown in Fig. 1-9. The waves proceeding within the horn continue on into space without changing their general direction of travel. At the aperture of the horn the phase front, *i.e.*, the surface along which the electric vectors have the same phase at a given instant, is a curved surface somewhat like the

cover of a rectangular watch with the convex side toward free space. By Huygens' principle, as used in elementary optics, each point along the phase front may be regarded as a new source of radiation. If the curvature of the phase front is slight, which is actually the case when the



FIG. 1-14—Illustrates vector addition of the fields arriving from elementary Huygens' sources. The angles between the successive vectors are phase angles. There is an infinite number of individual Huygens' sources but an approximate result may be obtained by dividing them into four or more groups.

taper of the horn is small, the phase surface approximates a rectangular section of a plane. At a distant point along the axis of the horn the waves arriving from the individual Huygens' sources are in the same phase, so that the amplitudes of the individual electric vectors add to make a large total. At a distant point along a line such as line CP_2 , at an angle to the axis of the horn, the waves arriving from the individual sources have different phases, the electric-field vectors adding vectorially as shown in Fig. 1-14. In this case the resultant field is only a fraction of that existing along the horn axis. It is clear that the radiator of Fig. 1-9 does not send electromagnetic waves equally in all directions but distributes them in space, in a manner illustrated in Fig. 1-15, in which angle Ψ is the angle between the axis of the horn and a direction in a plane passing through the axis of the horn. The radius vector drawn from zero to a point on the locus C is proportional to the electric field along the surface of a large sphere having its center at the center of the aperture of the horn.

The radiator of Fig. 1-1 results in a somewhat different distribution of electric field in space. We can again get a qualitative idea of the distribution of the radiated field by applying the elementary version of Huygens' principle. In this case the phase front at the aperture of the radiator is a curved surface which in the first rough approximation is a portion of a cone, as shown in Fig. 1-16. Because of the symmetry about the axis OZ of the antenna, the electric field has the same value along any small circle such as C on a large sphere S . The electric field, however, varies from point to point along a great circle passing through Z . In a direction ON , which is normal to the conical phase front, the waves originating from the individual Huygens' sources on the phase front are more nearly in phase than in other directions such

the waves arriving from the individual sources have different phases, the electric-field vectors adding vectorially as shown in Fig. 1-14. In this case the resultant field is only a fraction of that existing along the horn axis. It is clear that the radiator of Fig. 1-9 does not send electromagnetic waves equally in all directions but distributes them in space, in a manner illustrated in Fig. 1-15, in which angle Ψ is the angle between the axis of the horn and a direction in a plane passing through the axis of the horn. The radius vector drawn from zero to a point on the locus C is proportional to the electric field along the surface of a large sphere having its center at the center of the aperture of the horn.

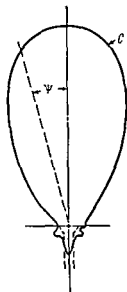


FIG. 1-15—A typical radiation pattern of a small horn antenna.

as OM . The result is that the electric field along a great circle (*e.g.*, one in the plane of the paper) is distributed as shown in Fig. 1-17. It is clear that the radiator of Fig. 1-1, like the horn of Fig. 1-9, does not radiate equally in all directions but has distinctive directional characteristics.

Because there are some planes, such as the planes of the small circles, within which the electric vector is constant in all directions, the radiator of Fig. 1-1 is sometimes called *omnidirectional*. The word *omnidirectional* as used in this connection is misleading because it applies only in a narrow sense.

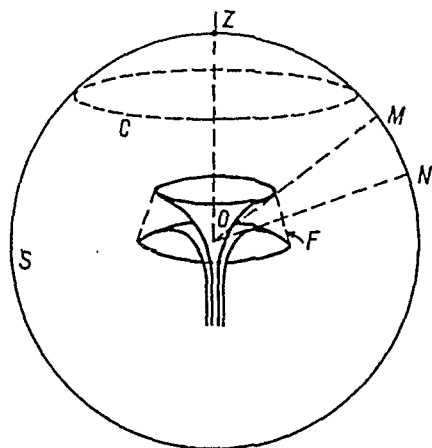


FIG. 1-16.—The phase front at the aperture of the radiator of Fig. 1-1; roughly conical as shown by broken line F .

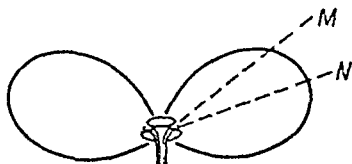


FIG. 1-17.—The main features of the radiation pattern of the antenna of Fig. 1-1.

1-7. Equivalence Principle and Radiation Patterns of Apertures.—

The application of the elementary form of Huygens' principle to antennas, such as the horn of Fig. 1-9 or the radiator of Fig. 1-1, is not justified when quantitative results are desired. A modification of this principle designed to take account of line charges along the edges of the aperture as well as of the vector character of the electromagnetic field is based on the so-called *equivalence principle*. In accordance with this principle the aperture is closed by an imaginary surface S carrying a distribution of electric charges, electric currents, magnetic charges, and magnetic currents which, taken together, exactly cancel the electric and magnetic fields of the aperture just outside the surface S . The field of the magnetic and electric currents and charges is calculated by the vector-scalar-potential method. The radiation field of the aperture is equal to this calculated field except for a change in sign. This procedure is accurate when the surface S is either a closed surface or a surface extending to infinity. Approximate results are obtained when the surface is extended only over the aperture, provided that the field outside the aperture is sufficiently small so that its contribution may be neglected.

The equivalent current and charge sheet along surface S has the following properties. The surface density of the distributed electric charge is such that it cancels the normal component ϵ_n of the electric field on one side of surface S . The electric current density i is such that

the equation of continuity is satisfied along S . This current density is that required to cancel \mathcal{H}_p , the component of the magnetic field \mathcal{H} parallel to surface S . Similarly, the magnetic charge density is chosen so as to cancel the normal component \mathcal{H}_n of the magnetic field on one side of surface S . The magnetic current density i^* , which is related to magnetic charge density σ^* by the equation of continuity, turns out to be just sufficient to cancel \mathcal{E}_p , the component of \mathcal{E} parallel to surface S . The direction of i is normal to \mathcal{H}_p , and the direction of i^* is normal to \mathcal{E}_p .

The magnetic and electric fields produced by the magnetic and electric charges and currents, after reversal in signs, are

$$\left. \begin{aligned} \mathcal{H} &= \text{curl } A - \frac{j\omega}{c} A^* + \frac{c}{j\omega} \text{grad} (\text{div } A^*) \\ &\text{in electromagnetic units} \\ \mathcal{E} &= -\frac{j\omega}{c} A + \frac{c}{j\omega} \text{grad} (\text{div } A) - \text{curl } A^* \\ &\text{in electrostatic units} \end{aligned} \right\} \quad (1-23)$$

where $A = \int \int \frac{[i]}{r} ds$, $A^* = c \int \int \frac{[i^*]}{r} ds$ and c is the velocity of propagation in free space. The magnitude of $i = (c/2\pi)\mathcal{H}_p$ and i is perpendicular to \mathcal{H}_p . The magnitude of $i^* = (1/2\pi c)\mathcal{E}_p$ and i^* is perpendicular to \mathcal{E}_p .¹

This method may be applied to calculate the field from the narrow slot antenna of Fig. 1-11. In this case it is convenient to take as surface S a plane parallel to the metal sheet into which the slot is built. If surface S is at a very short distance from the metal, the following assumptions are approximately correct. Within the aperture the normal component \mathcal{E}_n of the electric field is zero. Outside the aperture, in the immediate neighborhood of the metal sheet, the parallel component \mathcal{E}_p of the electric field is zero. Since $i = (c/2\pi)\mathcal{H}_p$ and $i^* = (1/2\pi c)\mathcal{E}_p$, it follows that $i^* = 0$ along the entire surface except within the aperture. Therefore the integral

$$A^* = c \int \int \frac{[i^*]}{r} ds$$

is extended only over the aperture. The electric current sheet along surface S contributes no field because it has an equal and opposite image in the metal sheet. When surface S is sufficiently close to the metal, the field of the image of the electric current in the metal sheet cancels the field of the current. This is true even of the current along that portion of the surface which is within the aperture, provided that the

¹ A more detailed discussion and references are given in ALFORD, ANDREW, *Ultra-short Electromagnetic Waves: V, Radiation, Elec. Eng.*, 62, August, 1943. This subject is related to recent work of Booker in connection with Babinet's principle.

aperture is narrow, for the direction of i is parallel to the long sides of the aperture. The field of the electric charges is also canceled out by the image.

The image of a magnetic current in a metal sheet is positive because the electric lines of force around a magnetic current are closed curves, as shown in Fig. 1-11*b*, so that two parallel magnetic currents flowing in the same direction result in cancellation of the tangential component of the electric field along the plane that is equidistant from the currents. The cancellation of the tangential component of the electric field is the condition that an image must fulfill.

The image of a magnetic charge is also positive because along a metal surface the normal component of the magnetic field must be zero, and this requirement is consistent only with a positive image.

As the distance between surface S and the metal sheet is decreased to zero, the field due to the electric currents and charges is canceled out by the images, while the field of the magnetic currents and charges is doubled. Thus we arrive at the conclusion that the field of a narrow slot in an infinite metal sheet is given by

$$\mathcal{H} = 2 \left[-\frac{j\omega}{c} A^* + \frac{c}{j\omega} \text{grad} (\text{div } A^*) \right] \quad (1-24)$$

$$\mathcal{E} = -2 \text{curl } A^*$$

where

$$A^* = c \int_s \int \frac{[i^*]}{r} ds$$

the integral being extended over the aperture of the slot and $i^* = \frac{1}{2\pi c} \mathcal{E}_p$ with the direction of i^* perpendicular to \mathcal{E}_p .

1-8. Elementary Dipole; Vector-scalar-potential Method; Radiation Patterns of Thin Antennas.—Another, alternate, approach to the problem of the distribution of field produced by a radiator of a given shape is based on the properties of an infinitesimal dipole. An exact solution for the distribution of electric field radiated from an infinitesimal dipole was obtained by Heinrich Hertz.¹ At a large distance from the dipole the electric field is distributed as shown in Fig. 1-18.

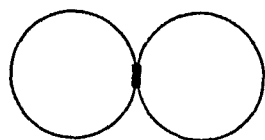


FIG. 1-18.—Radiation pattern of a small vertical dipole in the vertical plane.

The utility of the dipole concept comes from the fact that thin radiators, such as the one illustrated in Fig. 1-19, may be thought to be broken into a number of small segments, each of which is equivalent to an

¹ PIERCE, G. W., "Electric Oscillations and Electric Waves," McGraw-Hill Book Company, Inc., New York, 1920; SLATER, J. C., "Microwave Transmission," McGraw-Hill Book Company, Inc., New York, 1942.

infinitesimal dipole. Consider the current distribution in a thin radiator shown by curve *I* in Fig. 1-19. If the radiator of Fig. 1-19 is assumed to be subdivided into small segments, each segment may be regarded as a dipole with a different current flowing through it. The current in a particular dipole located at a distance z from the end of the radiator is a function of z . The charges on the ends of the adjacent dipoles do not quite cancel out. The difference between these charges, however, comes out to be exactly equal to the surface charges required by the conditions of continuity. Once a radiator has been broken up into dipoles, it is easy to visualize at least the main features of the radiation pattern with-

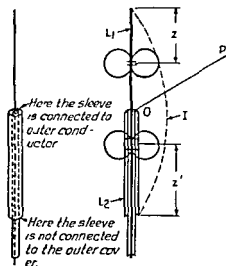


FIG. 1-19.—One form of sleeve antenna. The dotted curve *I* shows the approximate form of the current distribution. The two figure-of-eight patterns are radiation patterns of two of the large number of dipoles into which the antenna may be subdivided.

it is clear that the maximum radiation is at right angles to L because in that direction the individual dipoles contribute electric vectors that are in the same phase at a large distance. Also, no radiation will be emitted in the direction of the axis of the antenna, because the individual dipoles radiate nothing in this direction. In an intermediate direction such as *OP* the total electric field is less than that at right angles to L for two reasons: (1) because each individual dipole radiates less in this direction and (2) because the electric vectors from the individual dipoles will not add arithmetically but will add vectorially, each having a different time phase.

out a detailed calculation. Consider the radiator of Fig. 1-19. Since both conductors are assumed to be thin, the reflection of the waves at the ends of the radiators is nearly complete.¹ Under these conditions there is a standing wave along L_1 , the current being distributed approximately as shown by curve *I* in Fig. 1-19. The current along L_2 is also in the form of a standing wave which is a continuation of the standing wave along L_1 when both L_1 and L_2 are $\frac{1}{4}$ wavelength long. The current along L_2 flows in the same direction as the current in L_1 and, further, the currents flowing at all points along L_1 and L_2 are in substantially the same phase. If L_1 and L_2 are broken into dipoles as shown in Fig. 1-19,

¹ This is only approximately true at the lower end of the sleeve in Fig. 1-19.

1-9. Radiation Patterns of Cylinders.—The radiator shown in Fig. 1-2c differs from that of Fig. 1-2f in that the radiating element is a fat cylinder. The reflection at the end of a fat cylinder is only partial. The reflected wave along the cylinder is not equal to the direct wave traveling outward, so that the phase of the total current along the radiator is no longer constant but varies progressively from point to point. The fatter the cylinder, the smaller is the reflected wave in comparison with the direct wave. For this reason radiation patterns of fat cylinders, at least in a qualitative way, tend to approach the radiation patterns of conductors carrying only waves traveling in one direction. In order to gain a picture of what happens under such idealized conditions, we shall assume that waves along conductor W in Fig. 1-20 are propagated from left to right

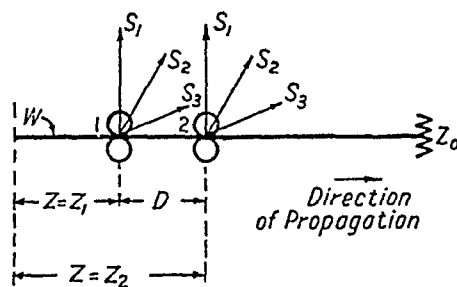


FIG. 1-20.—A diagram used in the discussion of a radiation pattern of a wire carrying waves that travel from left to right. The source of energy is not shown. The terminating impedance Z_0 , shown diagrammatically, could be a resistor in series with an inductance and a small spherical capacitor.

and that they are not reflected at the end of the conductor. If conductor W is broken into elementary dipoles, the phase of the current in each dipole increases linearly with z . Consider the radiation from dipoles 1 and 2, separated by a distance D . The phase of the current in dipole 2 is retarded with respect to the phase of current in dipole 1 by the number of electrical degrees corresponding to distance D . In the direction such as S_1 each dipole radiates maximum electric field, but the two electric vectors differ in time phase with respect to each other by the angle $2\pi D/\lambda$ and the fields add vectorially. In a direction such as S_2 each dipole radiates a somewhat weaker electric field, but the two fields add more nearly in phase because the radiation from dipole 1 has to travel a longer distance, and this fact reduces the phase advance of the field from this dipole with respect to the field from dipole 2. In a direction such as S_3 the fields from dipoles 1 and 2 are very nearly in phase, but the individual fields are very small. Similar arguments could be extended to apply to any number of dipoles along conductor W . When this is done, it is found that the maximum field results in some direction such as S_2 which is intermediate between the direction of the conductor and the direction at right angles to the conductor. The exact value of the angle between

the conductor and the direction of maximum radiation depends on the length of the conductor in wavelengths. The radiation pattern of a conductor carrying traveling waves is thus tilted toward the direction of propagation.

It is possible to carry out an accurate calculation of such patterns for conductors of various lengths when the conductor is assumed to have a diameter small in comparison with the wavelength. The results of such a calculation are illustrated in Fig. 1-21.

Since even along very fat cylinders the amplitude of the reflected wave is not zero but has a finite value, the radiation patterns of such antennas are intermediate between the patterns for traveling waves, illustrated in Fig. 1-21, and the well-known patterns of conductors carrying standing waves.



FIG. 1-21.—(a) Radiation pattern of a thin terminated wire two wavelengths long; (b) radiation pattern of a thin wire four wavelengths long.

When the distribution of the amplitude and phase of the current along a conductor or conductors is known at every point along their surfaces, the radiation field may be calculated by subdividing the conductor into elementary dipoles or by the equivalent procedure that makes use of the scalar and vector potentials. Both methods are well known, so that no derivation need be given here. The vector-scalar-potential method, when applied to u-h-f currents, gives the magnetic and electric fields in the following form:

$$\left. \begin{aligned} \mathcal{E} &= -\frac{j\omega}{c} \mathbf{A} + \frac{c}{j\omega} \text{grad} (\text{div } \mathbf{A}) \\ \mathcal{H} &= \text{curl } \mathbf{A} \end{aligned} \right\} \quad (1-25)$$

where

$$\mathbf{A} = \frac{1}{c} \int \int \frac{[\mathbf{i}]}{r} ds$$

and the integration is carried out over the surface of the conductor.

1-10. Relation between Radiation Patterns of Conductors and Apertures.—It is instructive at this point to compare the radiation pattern of a conductor with the radiation pattern of an aperture in an infinite plane.

Accordingly, we shall compare Eq. (1-24) of Sec. 1-7 with Eq. (1-25). Such a comparison shows that there is only one important difference: \mathcal{H} and \mathcal{E} are interchanged. The distribution of field due to an aperture in an infinite flat sheet is thus identical, except for the interchange of \mathcal{H} and \mathcal{E} , with that due to a flat conductor, of the same shape as the aperture, carrying current that at every point is in the plane of the conductor, normal to the electric field in the aperture, E_p , and proportional to E_p . This similarity of radiation fields of apertures and conductors, of course, holds good only on that side of the infinite plane where the aperture produces a field. For example, the field of a narrow aperture a half wavelength long (in which \mathcal{E}_p is at right angles to the long dimension) is identical with that of a half-wave antenna except that the lines of magnetic force are replaced with lines of electric force, and vice versa.

The lines of electric force due to an aperture are semicircles terminating on the metal sheet. The radiation pattern in a plane at right angles to the long dimension of the aperture is a semicircle. The radiation pattern in a plane through the long axis of the aperture is given by $\frac{\cos [(\pi/2) \cos \theta]}{\sin \theta}$. This is also the well-known pattern of a thin half-wave antenna.

1-11. Some General Properties of Directional Antennas.—Many applications call for broad-band antennas with highly directional characteristics. The property that makes an antenna very directional has been mentioned in connection with the horn of Fig. 1-9. This property, stated in a general way, is that the phase front in the immediate neighborhood of the antenna must be a flat, or at least approximately flat, surface that has substantial dimensions in comparison with the wavelength. The horn of Fig. 1-9 is a good example of a broad-band directional antenna. The horn produces a nearly flat phase front. Substantial changes in frequency do not result in substantial changes in the shape of the phase front near the aperture of the horn. The horn therefore remains very directional as the frequency is varied. As the frequency is increased, the area of the aperture in terms of square wavelengths increases and so does the directivity.

It is customary to measure directivity in terms of *beam width*, which is defined as that angle within which the electric field does not fall below 0.707 of its maximum value. Under certain conditions, which usually are approximately satisfied in practice, there is a simple relation between the beam width and the dimensions of an aperture. Consider a rectangular aperture with a horizontal dimension a and vertical dimension b . Assume, further, that the phase front is a plane and that the intensity within the phase front is uniform. If the aperture is divided into two equal parts by a vertical line through the center, there are two

apertures which act as two sources in the direction normal to the aperture. In a direction at an angle θ from the normal, in the horizontal plane, the waves coming from the two apertures are no longer in phase. The average phase of the waves coming from the left aperture differs from the average phase of the waves coming from the right aperture by the phase angle which is equal to $(2\pi/\lambda)(a/2) \sin \theta$. When this phase difference is 90 deg, the sum of the electric fields from the two apertures is 0.707 of the maximum value along the normal. This occurs when angle θ is such that

$$\sin \theta = \frac{1}{2} \frac{\lambda}{a}$$

When angle θ is small, $\sin \theta \approx \theta$ in radians, so that the beam width $W = 2\theta$ is

$$W = 2\theta = \frac{\lambda}{a} \text{ rad} = 57.3 \frac{\lambda}{a} \text{ deg}$$

This calculation disregards the directivity of the radiation from each half of the aperture. When this directivity is taken into account the result is

$$W = 2\theta = 51 \frac{\lambda}{a} \text{ deg.}$$

A convex phase front results in a broader beam than a uniformly illuminated one. When the illumination is concentrated toward the center of the aperture and falls off toward the edges, the beam width is increased.

When the illumination is not uniform, it is still possible to write

$$W = K \frac{\lambda}{a}$$

in which K is a characteristic of a particular type of illumination.

In at least one direction a directional radiator produces greater field than a nondirectional one when both are energized with the same amount of power. By supplying two radiators of different directivities with different values of power it is possible to find that power ratio P_1/P_2 which causes equal fields to be produced by the two radiators at a distant point. The power ratio P_1/P_2 is called the *power gain* of one radiator with respect to the other.

It has been found convenient to give the power gain of directional radiators with respect to a so-called *isotropic radiator* which radiates *linearly polarized* waves equally in all directions. While such a radiator does not and cannot exist, it is useful as a standard of comparison. All values of power gain in this book are given with respect to the isotropic radiator except when otherwise stated.

The following simple relation between power gain and beam width is useful when making crude estimates. Let W_1 deg be the beam width of a given directional radiator in the horizontal plane and W_2 be the beam width of the same radiator in the vertical plane. The beam can then be described as occupying $W_1 \times W_2$ sq deg. The total number of square degrees in a sphere is approximately 41,000. The power gain G of the directional radiator with respect to an isotropic one is therefore given by the relation

$$G = \frac{41,000}{W_1 \times W_2} \quad (1-26)$$

This formula is not exact because it fails to take into account the exact shape of the beam and the minor lobes. In spite of these shortcomings this formula may be used in making estimates except when the beam has an unusual shape.

Consider a radiator with a rectangular aperture with sides a and b . Then

$$W_1 = K_1 \frac{\lambda}{a}, \quad W_2 = K_2 \frac{\lambda}{b}$$

When these values of W_1 and W_2 are substituted in Eq. (1-26), we get

$$G = \frac{41,000 \, ab}{K_1 K_2 \lambda^2} \quad (1-27)$$

This equation is a special case of a more general equation

$$G = A \frac{\text{area of aperture}}{\lambda^2} \quad (1-28)$$

When the aperture is large and is uniformly illuminated, $A = 12.56 = 4\pi$. For nonuniformly illuminated apertures the value of A is lower than 12.56. For example, for apertures of parabolic reflectors the value of A is usually between 6 and 8. When Eq. (1-28) is applied to small apertures, *i.e.*, apertures of the order of λ^2 , the results are not reliable because the approximate formula for the beam width used in deriving Eq. (1-28) is inaccurate when a is comparable with λ .

In the chapters that follow, various types of broad-band antennas are discussed in detail. A number of the ideas and mental pictures sketched in this introduction are applied to specific cases and sometimes are also considered from a different point of view. The art of broad-band antennas is still very young, with the greater part of the realm as yet unexplored. It is hoped that the data and ideas described and compiled in the following pages will stimulate interest and further efforts.

CHAPTER 2

ULTRAHIGH-FREQUENCY MEASUREMENTS

BY J. A. NELSON, D. LAZARUS, J. W. CHRISTENSEN, AND R. R. BUSS

2-1. Transmission-line Measurements.—Since discussions of the theory of transmission-line measurement are available in many standard texts, in this section the discussion is limited to the transmission-line-measurement techniques that are of particular importance in the development of broad-band antennas. The equipment and techniques described are applicable to measurements at frequencies from below 100 Mc to above 3000 Mc.

It is possible to determine the input impedance of an antenna (or other device) by measuring the voltage or current distribution on the transmission line feeding the antenna. If the ratio of the voltage maximum to the voltage minimum, the distance from the antenna to a voltage minimum, the characteristic impedance, and the velocity of propagation in the line are known, the antenna impedance may be determined, as shown in Sec. 2-14.

2-2. The Smith Chart.—The Smith chart¹ is of such great usefulness in the solution of impedance relations determined from transmission-line measurements that a description of the calculator and brief directions for its use are included here.² The calculator can also be used for the solution of waveguide transmission problems (Sec. 2-20).

The chart, Fig. 2-1, is essentially a transmission-line circle diagram, with all impedances normalized or converted to a per-unit basis by expressing them as fractions of the characteristic impedance Z_0 of the line, as indicated in Fig. 2-1. Admittance, in the form $Y = G + jB$, is similarly normalized by dividing actual components by the characteristic admittance Y_0 of the line. Admittance problems are then solved in exactly the same manner as impedance problems by interpreting the designations of *resistance component* and *reactance component* on the chart as *conduct-*

¹ SMITH, PHILLIP H., Transmission Line Calculator, *Electronics*, 17, 1 (1944); 12, 29 (1939). A plastic calculator of considerable utility is manufactured and marketed by the Emeloid Co., Arlington, N.J.

² A chart useful for similar purposes is described by H. Cafferata, *Marconi Rev.*, January-February, 1941, p. 12; March, 1941, p. 21. This chart gives impedance in the polar form. An enlarged chart, giving attenuation in decibels, rather than nepers, has been constructed for the U.S. Naval Research Laboratory and is published by the U. S. Hydrographic Office as *Publications H. O. Misc. 9999* and *9999a* (a smaller form)

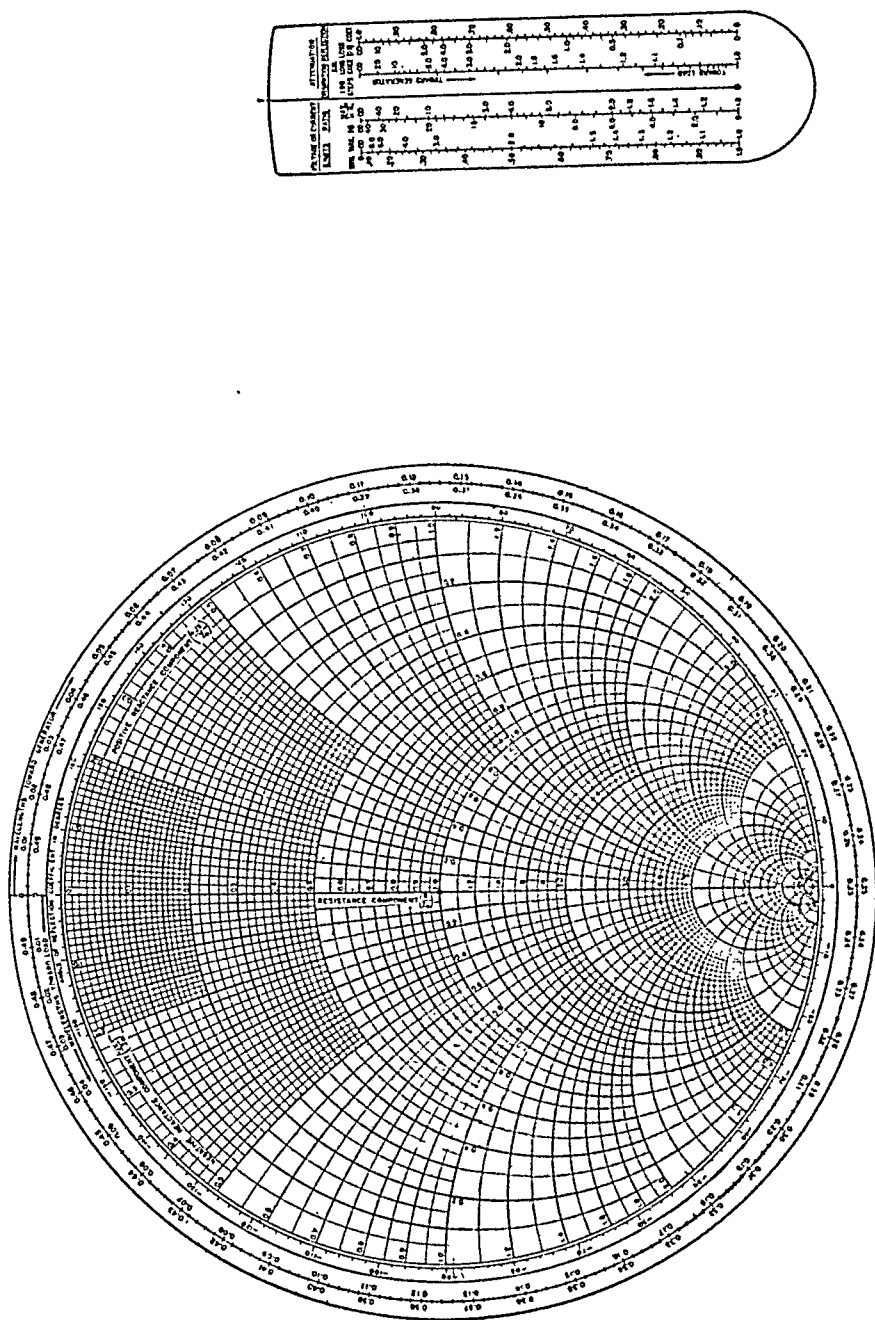


FIG. 2-1.—The Smith chart for impedance calculation. (By permission, from "Transmission Line Calculator" by Phillip H. Smith, Bell Telephone Laboratories, Electronics, January 1939 and January 1944.)

ance component and susceptance component, respectively (capacitive susceptance is positive).

The chart is composed of an orthogonal system of coordinate lines of constant series resistance and constant series reactance, all passing through the point of infinite impedance. The movable scales (the circular wavelength scale at the edge and the transparent arm and indicator) in conjunction with the chart relate the impedance (or admittance) at any point along a uniform transmission line to the impedance (or admittance) at any other point and to the several other electrical parameters such as voltage or current standing-wave ratio (σ), attenuation, and reflection loss. Circles of constant radius, centered at the point $R/Z_0 = 1$, $X/Z_0 = 0$ (the center of the chart), correspond to fixed values of standing-wave ratio. For example, the impedances anywhere on a uniform lossless transmission line having a voltage or current standing-wave ratio σ of 2.0:1 all lie on the circle passing through the points $R/Z_0 = 2.0$, $X/Z_0 = 0$; $R/Z_0 = 0.8$, $X/Z_0 = \pm 0.6$; $R/Z_0 = 0.5$, $X/Z_0 = 0$. Thus, since the σ of a lossless uniform line remains the same throughout its length, if either the impedance at any point on the line or the σ of the line and the position of a voltage maximum or minimum are known, the impedance at any point can be found by setting the slider on the transparent arm at the impedance or at the standing-wave ratio σ (which equals the per-unit resistance R/Z_0 at a voltage maximum) and rotating the arm the desired number of wavelengths toward the load or generator as required. Since impedances repeat every half wavelength, an integral number of half wavelengths are subtracted from the required wavelength distance to be traveled so as to bring the remainder to a quantity less than $\frac{1}{2}$ wavelength. Total attenuation may be introduced into the determination at any point of the calculation by moving the slider the required number of decibels toward the center or toward the rim of the chart, according to whether the unknown quantity to be determined is closer to the generator or to the load, respectively.

Reflection loss and the loss coefficient due to additional copper and dielectric loss that result when standing waves are on the line can be evaluated from the nomograms on the transparent arm, as well as such quantities as σ expressed in decibels (impedances disregarded), and maximum and minimum values of current and voltage (compared with values on a matched line carrying equal power).

Example: An antenna exhibited an impedance corresponding to a σ of 1.2 when measured with a 50-ohm slotted line. The minimum-voltage point shifted 0.13 wavelength toward the generator when the measuring equipment was short-circuited at the antenna. What is the antenna impedance?

Solution: Set the indicator at a σ of 1.2 and rotate the arm 0.13 wavelength

toward the generator from the zero-impedance end of the resistance axis since the low-resistance point on the line with the antenna connected is 0.13 wavelength farther from the generator than the reference zero-resistance (short-circuit) point. Then $Z/Z_0 = 1.0 + j0.19$; so $Z = 50 + j9.5$ ohms.

2-3. Components of Impedance-measuring Equipment for Wide Frequency Ranges.—The essential components of a transmission-line

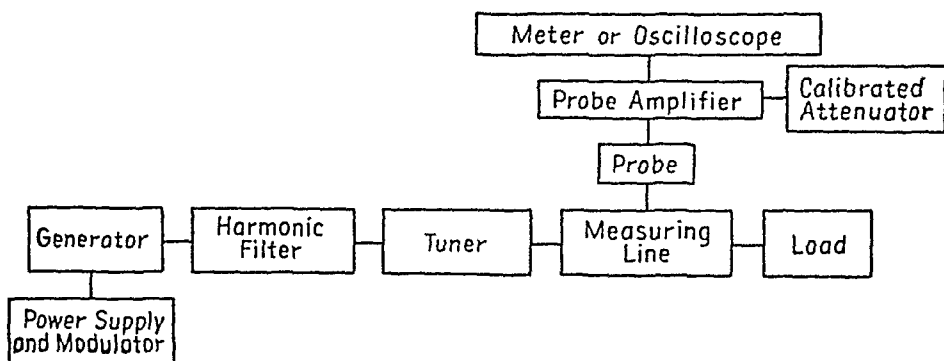


FIG. 2-2.—Block diagram of impedance-measuring equipment.

measuring setup are a slotted coaxial measuring line, a traveling probe, a probe-output indicator unit, and a signal generator (see block diagram Fig. 2-2). Each component should be designed to require a minimum of adjustment for frequency changes over a wide range. A measuring line with continuous dielectric support for the inner conductor (see Fig. 2-3) and having no discontinuities at the load end of the line may be used at any frequency for which the dielectric material has negligible loss and the line diameter is not large enough to permit the existence of modes other than the primary mode. The line must, of course, be at least $\frac{1}{2}$ wavelength long at the lowest frequency. A line long enough to use at 80 Mc is unnecessarily long for use at 3000 Mc. Continuous dielectric support for the inner conductor is desirable in a

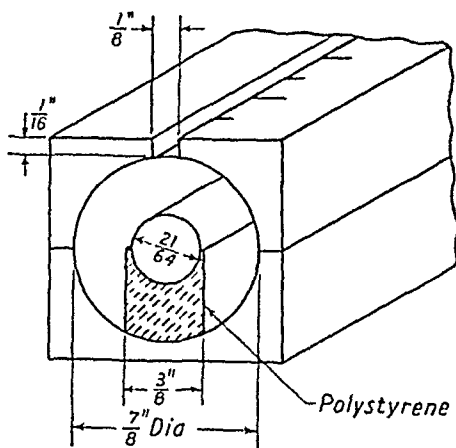


FIG. 2-3.—Typical construction of a 50-ohm measuring line.

measuring line for use at low frequencies, since a long inner conductor supported only at the ends is liable to sag or become bowed, and since dielectric-bead support distorts the voltage distribution on the transmission line. A measuring line for very high frequencies may be an air-dielectric line

with bead supports at both ends of the line. For these reasons it is desirable to have several measuring lines for different portions of the frequency range. It is sometimes desirable to have measuring lines of various diameters to eliminate the use of tapered connections (see Sec. 2-7).

The use of an untuned probe is recommended for broad-band work, since no time need be spent in tuning the probe, and the signal is at all times apparent on the indicator.

2-4. Detector Probes.—A probe consists of a shielded box that may be moved along a slotted measuring line and that has a voltage pickup wire or a current pickup loop extending through the slot into the field inside the measuring line (Fig. 2-4). The probe shown in Fig. 2-4a is an r-f output probe, with no detector. The probe in *b* is a crystal probe, in which one end of the pickup wire extends into the coaxial line, and in which the other is connected to two crystal detectors, both grounded on the far side, one directly and the other through a capacitor. The detected signal is the d-c voltage between the inner and outer conductors of the probe concentric output line. The probe in *c* is a bolometer probe, in which two bolometers are used. The r-f current in the pickup wire flows to ground through the two bolometers, which are in parallel for r-f currents since the capacitor acts as an r-f

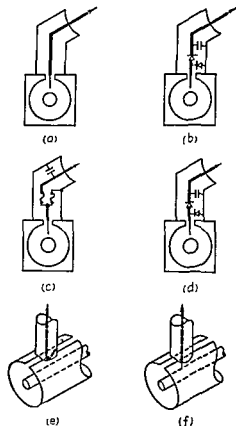


FIG. 2-4.—Various traveling-probe types (a) r-f output probe, (b) crystal probe, (c) bolometer probe; (d) diode probe, (e) current probe, (f) voltage probe.

short. The two bolometers are in series as far as the probe output is concerned. The probe in *d* is a diode-detector probe. At *e* and *f*, respectively, are shown a loop current probe and a pickup wire voltage probe. The spacing between the probe and the inner conductor should be kept as large as possible.

Usually the traveling probe on a measuring line has a detector such as a crystal, a diode, or a bolometer built into it. In general, the bolometer-type probe is preferable, since the bolometer is known to have

square-law response and requires no calibration. A crystal or diode must be calibrated frequently to ensure accuracy. The d-c output of the diode or crystal may be measured directly with a microammeter, or with a vacuum-tube voltmeter. When a crystal or diode is used, the u-h-f signal may be modulated or not, as is most convenient.

A bolometer is a small piece of fine wire, the resistance of which is proportional to temperature and, therefore, over the normal temperature ranges encountered, proportional to the square of the current flowing through the wire. The current in the wire is, of course, proportional to the voltage or current picked up on the transmission line by the pickup wire. A Littelfuse 8AG $\frac{1}{2}$ 00-amp fuse makes a satisfactory bolometer, although bolometers of smaller size and greater sensitivity are available for use at very high frequencies. It is customary to increase the sensitivity of a bolometer by applying a d-c voltage that heats the bolometer almost to the burn-out point. A modulated u-h-f signal is used on the transmission line in the conventional application of the bolometer. As a result of the modulation, the bolometer resistance varies at an a-f rate and gives rise to an a-f voltage, proportional to the square of the voltage on the transmission line, which may be impressed on the first stage of an audio amplifier. The amplifier output may be fed to a meter, or to an oscilloscope. When wide voltage variations, *i.e.*, large standing-wave ratios, are to be read, it is convenient to have a calibrated attenuator built into the circuit and to read the ratio of attenuator settings required to equalize the maximum and minimum signals. The input-to-output ratio of the amplifier must be constant for a wide input voltage range. A tuned amplifier may be used when the audio modulation is sinusoidal and of a fixed frequency.

2-5. Use of Probe with Receiver.—It is possible to use an r-f output probe with a receiver for an indicator unit. Such a system is very useful when the generator power is small. It is necessary either to calibrate the receiver or to use a calibrated r-f attenuator in either the probe or the generator circuit. The attenuator is used to set the receiver reading to the same value when the probe is set at the maximum and minimum voltage points on the slotted line. The ratio of attenuator settings is the standing-wave ratio. Two common sources of error in making standing-wave ratio measurements with an attenuator are (1) overloading of the receiver and (2) variable impedance at either end of the attenuator. The latter may be avoided by inserting some high-attenuation cable on *each side* of the attenuator. Usually the generator output is great enough to allow the use of sufficient lengths of lossy cable.

2-6. Calibration of Probes and System.—A convenient way to calibrate a detector, or to check the accuracy of any part of the equipment, is to measure the voltage distribution along the transmission line when

it is short-circuited at the load end. Any deviation from the expected sine curve indicates an inaccuracy of some sort. For instance, a bump in the curve indicates some mechanical inaccuracy, usually in the relative position of the inner conductor with respect to the outer conductor.

A small displacement of the inner conductor of the transmission line causes considerable error in the probe readings, since the field strength at the pickup wire is inversely proportional to the distance between the probe and the center of the inner conductor. The inner conductor may be displaced quite easily if the dielectric supports swell or are distorted in any way, or if the inner conductor becomes bowed or bent. The effect of such displacements on the standing waves on the line, however, is small, because seldom is the inner conductor very much out of alignment, and the characteristic impedance changes slowly as the line becomes eccentric.¹ The transition from the nominal characteristic impedance to the characteristic impedance of the eccentric line is gradual. For these reasons, the actual standing waves are distorted very little, although the probe readings may be badly distorted at the position of the displacement. It is possible, therefore, to make fairly accurate measurements on a measuring line with "bumps" when the probe readings may be taken at positions away from the bumps.

2-7. Line-to-load Connections.—One of the problems of transmission-line impedance measurements is the connection between the measuring line and the load to be measured. Seldom are the dimensions such that the inner and outer conductors of the load may be attached directly to the inner and outer conductors of the measuring line. Therefore it is usually necessary to devise some sort of tapered connection between the measuring line and the load. Sometimes it is necessary to employ connector adapters to connect the end of the taper to the connector on the antenna. Such tapers and adapters, when designed to maintain the specified characteristic impedance throughout, are not much of a problem at low frequencies. For instance, at 200 Mc there was no significant difference between the measured standing-wave ratio of a given load with

¹ The formula for the characteristic impedance of a concentric line is

$$Z_0 = 138 \log_{10} \frac{b}{a}$$

and of an eccentric line is

$$Z_0 = 138 \log_{10} \frac{b^2 - c^2}{ab}$$

where a and b are the radii of the inner and outer conductor, respectively, and c is the distance between the two centers. The difference in the two formulas is $(-c^2/ab)$, which is much smaller than b/a for reasonably small values of c . For instance, when $a = 0.125$, $b = 0.312$, and $c = 0.062$, Z_0 is 55 ohms for the concentric line and 53 ohms for the eccentric line.

one pair of connecting type-N connectors and with 10 connecting adapters of various sorts. At frequencies over 1000 Mc, tapers and adapters may have a large effect. In general, the best policy is to use as few devices as possible between the measuring line and the load. Figure 2-5 shows the standing-wave-ratio variation with frequency introduced by a type-N connector in a 50-ohm line, the standing-wave ratio of which is close to 1.05.

In Sec. 1-5 the effect on the standing-wave ratio of two sources of discontinuity along a transmission line is discussed. It is shown there that the highest standing-wave ratio is the product of the standing-wave ratios due to each discontinuity and the lowest is the quotient. When there are more than two discontinuities, the lowest standing-wave ratio is the combination of products and quotients that gives the lowest value greater than unity. Thus, four connectors inserted at random in a matched

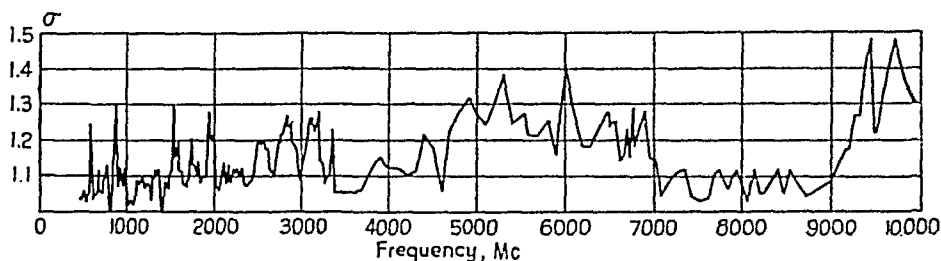


FIG. 2-5.—Standing-wave ratio produced by the insertion of a type-N connector in a 50-ohm load line.

line, each causing a standing-wave ratio of 1.15, will raise the standing-wave ratio to $(1.15)^4$ or 1.75 in the worst case, and in the best case the effects will completely cancel.

Occasionally one finds u-h-f types of connectors on equipment. These connectors introduce both loss and standing-wave-ratio distortion at frequencies in the decimeter range. Although the effects are small at 100 or 150 Mc, it is wise whenever possible to avoid their use with measuring equipment.

For many purposes it is desirable to have a dummy load that matches the measuring line. Such a load was used in measuring the effect of the type-N connector in Fig. 2-5. The problem of making a very broadband 50-ohm load has not yet been solved. A roll of 50-ohm r-f cable makes a fairly good load, having a standing-wave ratio of around 1.1:1 up to 10,000 Mc. A shorter length of lossy cable, such as RG-21/U, may be used. Disk resistors are available for low-power loads, but the resistance is not always constant and there is a reactive component in the impedance of most disk resistors. Loads consisting of carefully dimensioned Polyiron plugs placed in front of the short-circuited end of a

designed to feed a 50-ohm coaxial line, so that the maximum power is transferred when the generator "sees" a 50-ohm line. When the generator has an excess of power, it is possible to insert a length of lossy cable between the generator and the measuring line. The impedance looking into the lossy cable will be close to 50 ohms no matter what the load impedance is, and consequently tuning devices are not needed.

2-11. Elimination of Harmonics.—The output signals of practically all generators in the u-h-f region have appreciable harmonic content. Harmonics existing on the transmission line distort the voltage distribution along the transmission line, and the standing-wave ratios measured by an untuned probe are thus false. It is possible to eliminate harmonics of the measurement frequency by using a low-pass filter (see Chap. 27) designed to cut off at a frequency below the second harmonic. A series of low-pass filters with overlapping frequency ranges is very useful in wide-range impedance measurements.

2-12. Ground Plane.—A great many antennas are designed to work against a ground plane, such as the metallic fuselage of an airplane. A laboratory ground plane should be as large as is convenient, preferably a full wavelength in diameter, and at least a half wavelength in diameter, at the lowest frequency at which it is to be used. It is possible to design a smaller ground plane by introducing lossy material, such as resistance cloth, around the periphery in such a way as to absorb the current flowing on the ground plane and prevent reflections either from the edge or from the lossy material. A good position for a ground plane is on the outside of a building, covering one wall, with an aperture in which the antenna to be measured may be placed. Precautions must be taken to ensure good contact from the base of the antenna to the ground plane. An added advantage of having the ground plane outside the building is that no difficulties are encountered with measurement errors due to reflections from objects in the radiation field.

2-13. A Complete Measurement Setup.—A block diagram of the measurement setup used at the Radio Research Laboratory is shown in Fig. 2-2. Two sets of equipment were used, one for measurements at frequencies below 1500 Mc and one for measurements at frequencies above 1000 Mc. In the high-frequency setup, the lines were shorter and the tuners smaller. The filters, of course, were designed to have appropriate cutoff frequencies.

2-14. Measurement Procedure.—A typical data sheet for an impedance measurement of an antenna is shown in Table 2-1. For these measurements, a heterodyne frequency meter was used on which the frequency was read directly. To locate the positions of minimum and maximum voltage, the traveling probe was moved along the transmission line, and the magnitude of the detected signal observed on an oscilloscope.

The voltage standing-wave ratio was determined by taking the square root of the ratio of attenuator settings required for the same oscilloscope level for the minimum and maximum voltages, since a bolometer probe was used. The Reference Minimum column lists the positions of voltage minimums measured with a short circuit at the load position. The distance d in centimeters is obtained by subtracting the position reading of the reference minimum from the position reading of the voltage minimum measured with the load on the line. When the latter is on the load side of the position of the reference minimum, the difference between the two is a negative distance, as in the $f = 280$ Mc row in Table 2-1. The distance d/λ is, of course, the distance d divided by the wavelength corresponding to the frequency. The pointer on a Smith chart (Sec. 2-2) is rotated the distance d/λ from the 0-wavelength point in a counterclockwise direction when d/λ is positive and in a clockwise direction when d/λ is negative. The antenna impedance is located on the pointer line

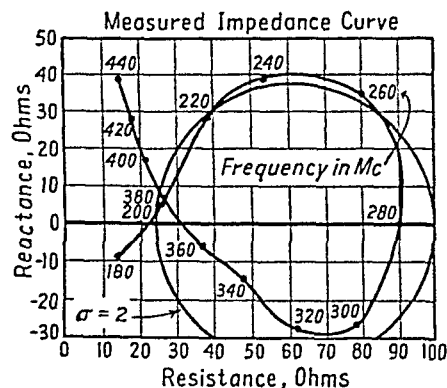


FIG. 2-8.—Measured impedance plotted on an R - X diagram.

In many cases only the magnitude of the reflection coefficient is needed and an impedance curve is not required. When the measuring line is of the same characteristic impedance as the feed cable for the antenna, the measured standing-wave ratio is the standing-wave ratio that will exist on the feed line. The coefficient of reflection ρ is related to the voltage standing-wave ratio σ by the formula

$$\rho = \frac{\sigma - 1}{\sigma + 1} \quad (2-2)$$

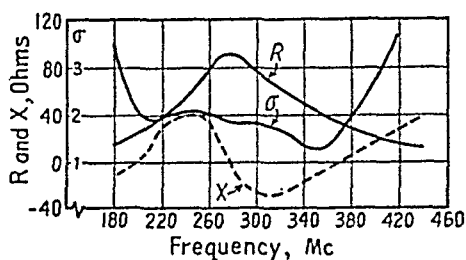


FIG. 2-7.—Measured impedance vs. frequency.

at the distance from the center corresponding to the measured standing-wave ratio σ . The characteristic impedance of the transmission line is 50 ohms, so that the values read off the Smith chart are multiplied by 50 to obtain the load impedance. The impedance curve may be plotted as in Fig. 2-7, resistance and reactance vs. frequency, or as in Fig. 2-8, resistance vs. reactance on an impedance plane, each point representing the impedance measured at a different frequency.

TABLE 2-1

Fre- quency	E_{min}^2	E_{max}^2	σ	E_{min} posi- tion	Ref. min. posi- tion	$\frac{d}{cm}$	$\frac{d}{\lambda}$	$\frac{R + jX}{50}$	$R + jX$
180	470	34	3.7	41.0	36.9	4.1	0.029	0.28 - j0.16	14 - j8
200	450	110	2.02	17.4	22.5	-5.1	-0.032	0.50 + j0.15	25 + j7
220	340	90	1.95	54.8	10.9	43.9	0.380	0.78 + j0.58	39 + j29
240	230	55	2.04	37.0	53.9	-16.9	-0.159	1.08 + j0.76	54 + j38
260	95	23	2.03	21.7	41.6	-19.9	-0.204	1.02 + j0.70	81 + j35
280	90	28	1.79	9.6	32.3	-22.7	-0.250	1.79 + j0	90 + j0
300	66	20	1.82	40.0	22.6	17.4	0.206	1.58 - j0.52	79 - j26
320	320	115	1.67	28.8	14.8	14.0	0.176	1.24 - j0.54	62 - j27
340	52	28	1.36	17.2	7.6	9.6	0.128	0.96 - j0.30	48 - j15
360	31	16	1.39	39.4	36.4	3.0	0.012	0.74 - j0.12	37 - j6
380	38	10.9	1.87	27.3	29.2	-1.9	-0.028	0.64 + j0.12	27 + j6
400	75	11.2	2.59	50.5	22.7	27.8	0.438	0.43 + j0.34	22 + j17
420	140	10.5	3.65	11.2	16.5	-5.3	-0.088	0.37 + j0.57	19 + j29
440	760	24	5.6	33.5	11.0	22.5	0.390	0.29 + j0.78	15 + j39

It is usually sufficient to plot the standing-wave-ratio curve, as the limits of acceptability of antenna impedance may be defined by the maximum acceptable standing-wave ratio.

2-15. Balanced-line Measurements Using Two Radio-frequency Cables.—Sometimes it is necessary to make measurements on a balanced

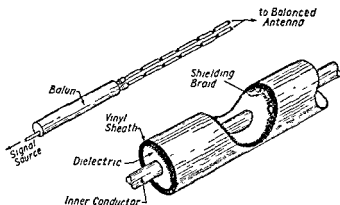


FIG. 2-9.—Balanced two-wire measuring system employing r-f cable. Enlarged view of cutaway.

line. One rather crude technique that is satisfactory and simple is to use a probe voltmeter and two lengths of r-f cable, fed by a balun (see Sec. 3-14) and cut open at regular intervals, as in Fig. 2-9. Such cuts do not change the characteristic impedance of the line appreciably. A probe voltmeter is a tuned voltage probe, where the inner conductor of the

probe makes contact with the inner conductor of the cable and the outer conductors of the probe and the cable are connected. The probe is so adjusted that it presents a very high impedance to the cable. A crystal detector and a microammeter are built into the probe voltmeter, and it is convenient to have a crystal calibration curve attached. The tuning device may be a small variable capacitor or a telescoping coaxial line, or a combination of both. It is usually necessary to use two probe voltmeters, since with one it is not possible to tell when the probe voltmeter is tuned to present a high impedance to the line. With two probe voltmeters, one is put on the line to serve as an indicator, and the other is tuned until there is a negligible difference in the readings of the first when the tuned probe voltmeter is either on or off the line.

2-16. Waveguide Measurements: Slotted Sections.—Measurements of reflection coefficient and impedance of a load or discontinuity in a waveguide transmission system may be made in a manner completely analogous to the methods used with coaxial transmission lines. Sections 2-16 to 2-22 are devoted to a description of some of the components and techniques employed in making measurements with rectangular waveguides in which energy is propagated in the fundamental or $TE_{1,0}$ mode.

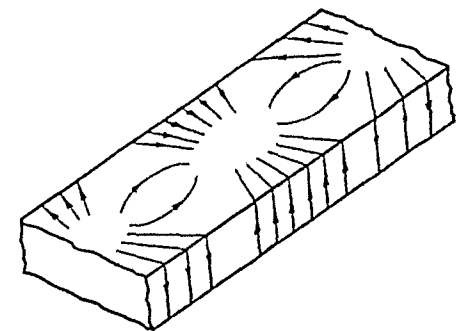


FIG. 2-10.—Lines of current flow in rectangular waveguide for $TE_{1,0}$ mode.

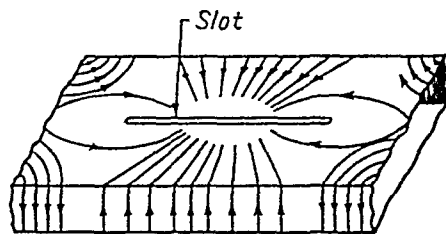


FIG. 2-11.—Proper position of slot in rectangular waveguide for $TE_{1,0}$ mode.

In a coaxial slotted section, the only restriction on the position of the slot milled in the outer conductor is that it be parallel to the axis of the line. Since the currents on the inside wall of the outer conductor travel parallel to the axis, the slot does not, therefore, disturb the field inside the line. In a waveguide, on the other hand, the electric field does not have radial

symmetry for the $TE_{1,0}$ mode, and care must be taken to locate the slot in such a position that it does not disturb the current flowing in the walls of the guide. The current distribution on the inner surface of the guide is obtained by taking the orthogonal trajectories of the magnetic field within the guide. This distribution is shown in Fig. 2-10. It is evident that only at the center of the broad dimension of the guide are the lines of current flow parallel to the direction of propagation. The slotted section must, therefore, be constructed with

the slot cut along this center line, as shown in Fig. 2-11. As in the coaxial slotted line, it is important also that the surface which is slotted be uniform, so that, as the probe is moved along the slotted section, it will always project the same distance into the interior of the guide.

2-17. Probes, Detectors, and Indicators.—For waveguide measurements, the same probes, detectors, and indicators are used as in measurements with conventional transmission lines. These components have been discussed in Sec. 2-4.

In the use of the probe, some consideration must be given to the effect of the penetration of the probe into the guide. Slater¹ has shown that the parallel impedance introduced into a waveguide by a probe is of the

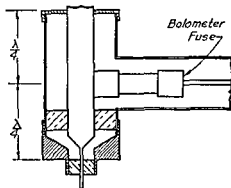


FIG. 2-12.—Resonant probe.

form $Z_p = Z/k$. The quantity Z is the sum of the intrinsic impedance of the probe plus the impedance looking from the probe into the line feeding it. The quantity k is a constant depending on the length of the probe relative to the smaller dimension of the guide and increases according to a square-law relation from a minimum of 0 to a maximum value of 2 with increasing length of the probe.

In order to maintain a large value for Z_p , either k must be very small or Z very large. The value of k can be kept small by keeping the length of the probe very short compared with the height of the guide. If test oscillators of low power output are used in the measuring setup, it is often difficult with such a limitation to extract sufficient power from the guide to ensure accurate measurements. It may then be better to make Z large by tuning the probe to resonance by means of a coaxial stub in series with the probe as in Fig. 2-12.

Figure 2-13 shows a typical slotted guide and probe assembly, designed for use at about 10,000 Mc. The probe is tunable to resonance, and the whole detector assembly is mounted in a massive carriage that may be moved along the axis of the guide by means of a rack-and-pinion drive. Although it is desirable for electrical reasons to keep the width of the slot milled in the slotted section to a minimum, this introduces mechanical problems at these very high frequencies, since the probe must be constructed to project through the slot. Therefore, it is necessary to make the slot slightly wider than might seem advisable from other

¹ SLATER, J. C., "Microwave Transmission," Chap. VII, McGraw-Hill Book Company, Inc., New York, 1942.

considerations. The amount of energy that might be lost by leakage through the slot is negligible, for practical purposes, but some error in reading may result if this r-f energy reaches the indicator circuit. For this reason, quarter-wave chokes are built into the underside of the carriage where it makes contact with the top surface of the guide. These chokes isolate the bolometer and the cable leading to the indicator from any r-f energy that might be radiated from the slot.

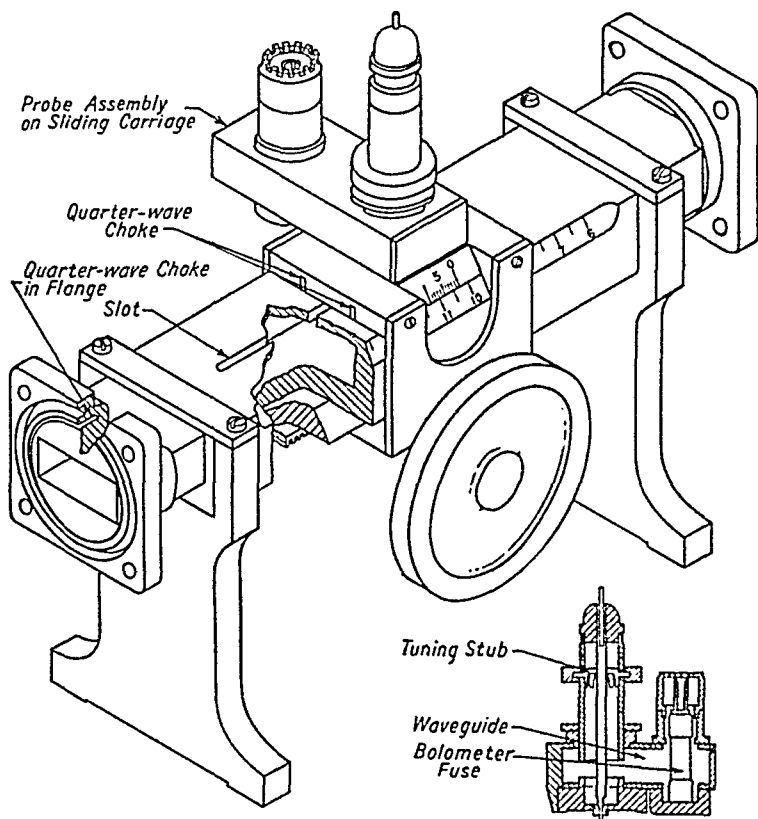


FIG. 2-13.—Typical probe and slotted-guide assembly.

2-18. Tuning Sections.—As in the case of coaxial tuning sections, a branching section of waveguide may be used as a tuning stub. If the stub branches from the narrow side of the guide as in Fig. 2-14a, the voltage and current at the junction are shared by the two branches. Such a stub is obviously of the shunt type and is identical in behavior to the coaxial tuning stub. If the stub branches from the broad side of the guide, as in Fig. 2-14b, the current in the direction of propagation is interrupted and flows into the branch circuit and then back into the main guide in a manner characteristic of a series circuit. In this series case, impedances add in the same way as admittances for the shunt case.

In either case, two or more stub sections may be used to effect a match to a given impedance.

Another very simple tuning mechanism commonly used in waveguide measurements is the so-called *tuning screw*. A screw or rod, inserted from the broad face of the guide, presents a shunt capacitance at that point. The value of the reactance can be varied by variation of the

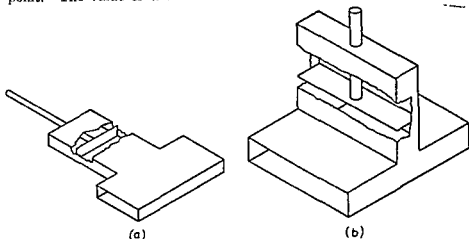


FIG. 2-14.—Waveguide tuning sections (a) shunt tuning stub, (b) series tuning stub. length of the screw. If the screw is also free to move back and forth in the direction of propagation over at least half a guide wavelength, as shown in Fig. 2-15, any impedance can be matched to the guide.

2-19. Waveguide Loads.—For certain types of measurements it is necessary to employ sections of waveguide that either absorb the energy traveling down the guide or radiate it into space. Such devices, known as *waveguide loads*, may exist in any one of a large variety of forms and are essentially comprised of an absorbing medium in combination with associated matching sections. By use of proper tuning sections it is possible to effect a perfect match either to space or to a dissipative

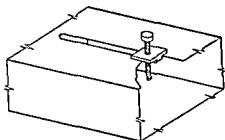


FIG. 2-15 —Tuning screw.

medium of any type at a given frequency. There is nothing in the design of a particular load, therefore, that makes it in any way unique. The various types of loads that are discussed here are chosen because of their relative simplicity and broad-band frequency characteristics.

Probably the simplest method of extracting energy from a waveguide without introducing large reflections is by use of a gradually tapered horn, as in Fig. 2-16a. If the angle of taper is small and the length and

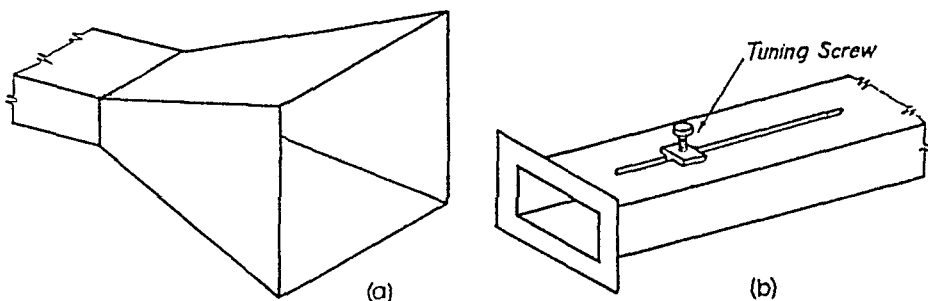


FIG. 2-16.—(a) Tapered horn; (b) guide terminated directly in air, with tuning screw.

dimensions of the aperture large compared with the wavelength, such a system will transfer all the energy from the guide to space without appreciable mismatch over a very broad band of frequencies. A similar transfer can, of course, also be effected simply by terminating the guide directly in air, as in Fig. 2-16b. Here some impedance mismatch is generally encountered and must be compensated by means of a tuning section before the aperture, as shown. The second method is naturally more frequency sensitive than the first. Both of these methods have the disadvantage that varying reflections may be introduced into the guide by objects moving about in the radiation field. For this reason, it is sometimes necessary to employ loads in which the energy may be completely absorbed without radiation.

Several types of such loads are shown in Fig. 2-17. In each of these a resistive medium such as carbon, or water, is used to absorb the energy. In each case the transition between the air-filled guide and the load is made as

gradual as possible to reduce the frequency sensitivity of the device, but for single-frequency use any reflections that may be introduced can always be canceled by the use of tuning sections.

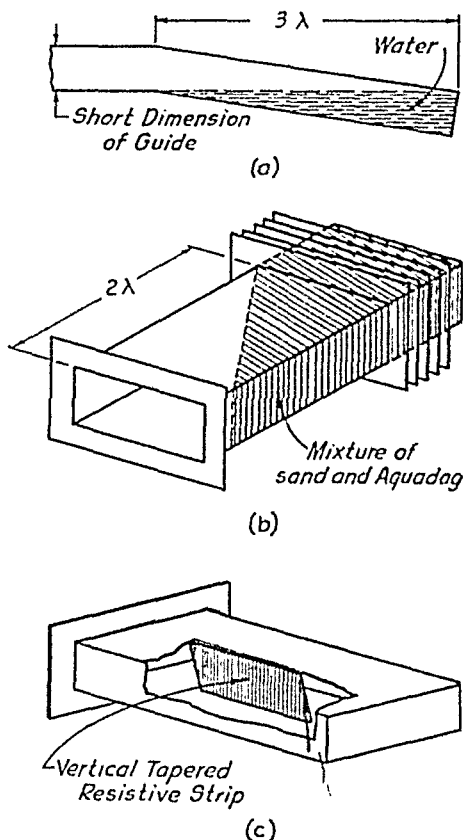


FIG. 2-17.—Absorption loads. (a) broad-band water load; (b) sand and aquadag load; (c) resistive strip load.

2-20. Measurement Procedure.—The techniques involved in measurements with waveguide transmission systems are, with minor exceptions, identical to those used in measurements with coaxial transmission lines. As has already been pointed out, the various waveguide components employed, slotted guide, tuning sections, loads, etc., are equivalent functionally to coaxial components.

Values of resistance and reactance obtained from standing-wave-ratio measurements in either coaxial line or waveguide are expressed quantitatively in terms of the characteristic impedance of the transmission line. In the case of coaxial lines, the characteristic impedance is a well-defined quantity that is dependent, for transmission in a given mode, only on the physical dimensions of the conductors making up the line. In waveguide measurements, some ambiguity arises from the possibility of defining three characteristic impedances for the guide, all of which are dependent on frequency. For measurements concerned with impedance matching, however, the closest analogy to coaxial transmission lines is obtained by employing the value of characteristic impedance defined by the ratio of the voltage between top and bottom surfaces of the guide to the current flowing in the walls of the guide, under the assumption that there is no reflected wave. For the $TE_{1,0}$ mode in air-filled waveguide, this characteristic impedance is given by the relation

$$Z_0 = \frac{\pi 377 b \lambda_g}{2a \lambda} \quad \text{ohms} \quad (2-3)$$

where a and b are the width and height of the waveguide, respectively, and λ_g and λ are the guide wavelength and free-space wavelength, respectively. This value of characteristic impedance differs only by a constant factor from that obtained by taking the ratio of the power flowing down the guide to the square of the current flowing in the walls

The impedance at any position along the waveguide is then equal to the ratio of the voltage to the current at that position. It is related to the impedance at any other portion of the guide by the conventional transmission-line equations, considering the phase-shift constant β to be given by

$$\beta = \frac{2\pi}{\lambda_g} \quad (2-4)$$

where λ_g is the wavelength in the guide and is given by the relation

$$\lambda_g = \frac{\lambda}{\sqrt{1 - \left(\frac{\lambda}{2a}\right)^2}} \quad (2-5)$$

With these modified definitions of impedance and phase shift, impedance measurements can be taken of a load or discontinuity in a waveguide

transmission system and the values of impedance determined and translated to other points in the system by the use of conventional transmission-line charts (see Sec. 2-2).

In broad-band antenna work, waveguide measurements are generally limited to two types:

1. Measurements of the impedance of a waveguide radiating element, such as a horn or slot, made in order to design appropriate matching sections
2. Measurements of the degree of mismatch introduced by bends or obstructions in an otherwise matched waveguide transmission system, without regard to actual values of impedance

To measure the impedance characteristic of a radiating element such as a horn or slot, a procedure identical to that followed in the measurement of the impedance of a coaxially fed antenna may be employed (Sec. 2-14). Standing-wave ratios are taken by means of a traveling probe and detector in a slotted section of guide and positions of the maximums or minimums determined. The radiating element is then replaced by a short circuit and the position of a maximum or minimum point of field intensity noted. Since the distance between this point and the short-circuited end of the guide must be an integral number of half guide-wavelengths, for a minimum value of electric field, the phase angle of the impedance of the radiating element is determined directly. With the standing-wave ratio and phase angle known, the impedance at the point at which the short circuit was inserted can be computed by means of a Smith chart (Sec. 2-2), or any of the other conventional methods. Numerical values of reactance and resistance may then be determined by computing the value of Z_0 by Eq. (2-3).

The general precautions that must be followed in making measurements of this sort are the same as those which apply in impedance measurements made with coaxial lines. The length of waveguide between the probing section and the radiating element must be kept to a minimum to minimize the error introduced by frequency shift of the signal generator. Care must be taken to provide filters for the signal generator in order to prevent harmonics of the fundamental frequency from being propagated in the guide. This is particularly important in waveguide measurements, since the harmonics will travel in the guide with phase velocities different from that of the fundamental, and the measured values of standing-wave ratio and phase angle will vary for different pairs of maximums and minimums, resulting in an appreciable error. The presence of harmonics in waveguide systems is easily recognized by the erratic nature of the standing-wave-ratio data. In general, the harmonics are less troublesome if the radiating element is also a broad-band radiator for the harmonic frequencies, since these waves will then be

transmitted without reflection. The major errors are encountered when the radiating element will transmit only the fundamental frequency and will reflect one or more of the harmonics.

In many cases impedance measurements are not required, but only a determination of the reflection coefficient introduced into a matched waveguide transmission system by an obstruction or other discontinuity. In these measurements, waveguide loads, such as those described in the preceding section, are employed. Two slotted sections are generally required, one between the generator and the discontinuity, and a second between the discontinuity and the load. This second section must be used to make certain that the waveguide is properly terminated by the load. The load must be adjusted to give unity standing-wave ratio as measured with the second slotted section. Then the reflection coefficient introduced by the discontinuity is determined directly by the value of the standing-wave ratio taken with the slotted section that is between the generator and the discontinuity.

2-21. Example of Waveguide Measurements.—A set of waveguide impedance measurements made in the manner described in Sec. 2-20 was taken to

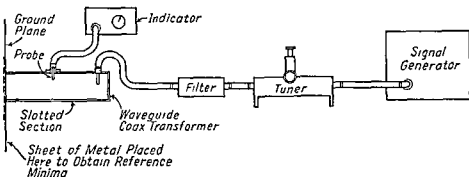


FIG. 2-18.—Experimental setup for waveguide measurements

determine the impedance at the aperture of a rectangular waveguide that was terminated directly in a ground plane. The experimental setup employed is shown in Fig. 2-18. It consisted of a signal generator, a set of filters, a waveguide-to-coaxial-line transformer section, a slotted section of guide with probe and indicator, a ground plane with an aperture cut in it equal to the cross section of the guide, and a sheet of metal to be placed across the aperture to serve as a short circuit. A cavity type of wavemeter was used to measure frequency. A plot of the impedance values obtained is given in Fig. 2-19.

2-22. Antenna Pattern Measurements. General Considerations.—Because of the large number of patterns needed to show the behavior of a broad-band antenna over its frequency range, determining the radiation characteristics of such antennas is laborious and unnecessarily time

consuming unless the techniques used are themselves broad band. They should be such as to permit the taking of data that conform to a given standard of accuracy throughout the frequency band with a minimum number of changes and adjustments of the equipment.

Basically, the techniques are the same as those used in any other phase of microwave antenna work. Advantage is almost always taken of the fact that the small physical size of the antennas permits a measuring process in which the antenna is rotated while the measuring equipment remains stationary. Obviously this is easier and faster than moving the test equipment around the antenna as is required at low frequencies. It is also, of course, perfectly simple to measure the pattern in any plane

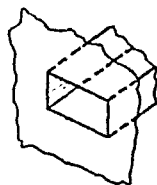
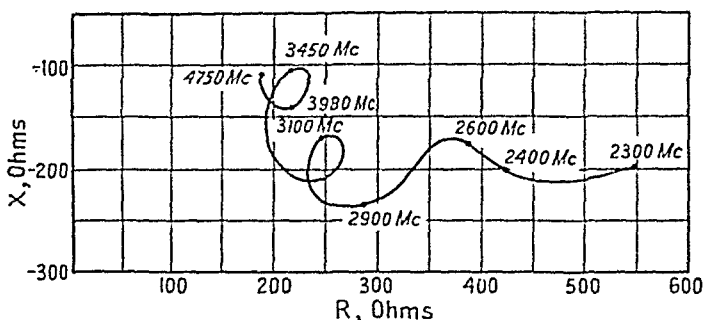


FIG. 2-19.— R - X plot of impedance at aperture of waveguide terminated in ground plane.

by appropriately orienting the antenna on the turning mechanism. Accordingly, the facilities that should be available for measuring the radiation patterns of broad-band antennas fall into the following categories:

1. Microwave signal sources that cover as nearly as practicable the frequency range of interest
2. Receivers or other detectors that also meet the general broad-band requirements
3. Broad-band test antennas to be used either to transmit to, or receive from, the antenna under test
4. Working site and mechanical facilities such as turntables and antenna mounts that will introduce either negligible reflections, or reflections whose behavior with frequency can easily be controlled
5. Auxiliary devices that speed the taking of data

In the following sections, typical specific techniques are indicated that have facilitated meeting the foregoing requirements. It is assumed that the reader is familiar with the basic considerations that apply to any radiation-pattern measurements. These would include such matters as the number of surfaces in which patterns should be measured in order to represent satisfactorily total three-dimensional patterns, the advantages of various coordinate systems, and the necessity for investigating the polarization as well as the intensity of the field being measured.

2-23. Basic Pattern-measuring Techniques.—It is well known that, in accordance with the Rayleigh-Carson reciprocity theorem, one may measure the radiation pattern of an antenna by using it either for transmitting or receiving. According to the first method, the relative field strength at a fixed distance from the antenna is measured in various directions with respect to the antenna. In the second method, the relative strength of the received signal is measured at a fixed distance from a transmitting antenna, as the direction of the transmitting antenna relative to the antenna under test is varied. Throughout the following discussion, it will be assumed in the interests of simplicity that the antenna under test is being used for receiving.

2-24. Signal Sources.—Almost all types of broad-band oscillators may be used in pattern measurements. The most useful are obviously those in which the greatest frequency range may be obtained with the greatest facility. Whether or not the signal should be modulated, and if so, the type of modulation most suitable depends on the nature of the oscillator and on the type of detector used in the receiving equipment. Such considerations naturally depend upon the particular circumstances. Practically all the commercial signal generators, however, and the oscillators (except those of the highest power) described in the receiver and transmitter sections of this book have been used without difficulty. There is one important precaution that may appear self-evident, but that is nevertheless mentioned because failure to observe it adequately has occasionally resulted in considerable nuisance. It consists of providing sufficient shielding and by-passing to ensure that negligible radiation occurs from either the oscillator itself or the cables leading from it. It is obviously important to shield the oscillator and its cables; otherwise the signal received will arrive not only from the transmitting antenna, but from the oscillator and the other objects connected with it. This is more likely to be troublesome in broad-band work than in other work, because shielding and filtering that are satisfactory over a great part of the range of an oscillator may be completely unsatisfactory over the remainder. Accordingly, a superficial test at a few frequencies, which indicates that oscillator radiation is negligible, can be misleading and result in erroneous data.

2-25. Detectors and Receivers.—Practically all that has been stated above in connection with signal sources applies as well to detectors and receivers. The effect upon the measurements of receiving equipment that responds to radiation which falls directly upon it or on the cables leading to it is likely to be just as serious as the effect of an oscillator that radiates. In both cases the magnitude of the effect depends upon the relative location of the pieces of the equipment and on the patterns of the antennas. But as a general rule, it is not possible to make satis-

factory pattern measurements if either the transmitting or the receiving equipment is not very well shielded. In this connection, it might be mentioned, attempts to compensate for poorly shielded equipment by placing it in shielded equipment houses are liable to be unsuccessful, because the larger the object being shielded (in this case the house), the more difficult the job is likely to be. Furthermore, such arrangements are likely to increase unnecessarily the seriousness of reflections at the measuring site and, therefore, can hardly be considered good practice.

Of the many types of receiving equipment suitable for pattern measurements, such as bolometers, crystals, and diodes (used with various types of amplifiers), and the superheterodyne and other receivers, no particular one is most suitable for all purposes. However, the circumstances of broad-band work frequently make the superheterodyne receiver much more useful in pattern measuring than it ordinarily is in microwave antenna work. When used for antenna measurements, the current in the second detector stage is ordinarily used as an index of received signal strength.

The superheterodyne receiver possesses several advantages. One of the most important of these is the ease and accuracy with which good receivers of this type can be tuned. A second is that gain can readily be controlled by varying the bias on the i-f amplifier stages. This permits the use of a variety of signal sources, some of which provide much more power than others, without producing either overloading with strong signals or excessive noise-to-signal ratios with weak ones. Because in broad-band work a heterogeneous collection of signal sources is likely to be unavoidable, the adjustable sensitivity can be an important point.

The disadvantages of the superheterodyne receiver are primarily that calibration curves must be used and that the amplitude range available at a given i-f a-g-c setting is relatively small. Neither of these is as serious in broad-band work as might be supposed. With regard to calibration curves this is true because, although a particular calibration curve applies only for the gain-control setting at which it was taken, such a curve is usually applicable regardless of frequency. The principal effect of the limited amplitude range of the superheterodyne receiver is that of decreasing the accuracy with which minor lobes of a radiation pattern can be measured. It is usually difficult to take readings that are more than about 25 db below the maximum reading. In broad-band work this frequently does not matter because no attempt at detailed control of minor lobes is made. It is inevitable that they vary with frequency, and it is usually sufficient that they be below a particular level. With the superheterodyne receiver it is usually possible to determine whether or not this is so.

2-26. Test Antennas.—The most satisfactory test antennas are those that are not only broad-band with respect to input impedance, but unidirectional and balanced as well. A unidirectional test antenna is practically always desirable in pattern measuring because it permits directing most of the radiation at the antenna under test. Consequently, a completely unobstructed test site is not needed. It is sufficient that there be no obstructions in the sector included in the major lobe of the transmitting antenna. Conversely, the amount of directivity needed will depend upon the site available. Frequently a clear 180-deg sector can be found, and under this circumstance a corner-reflector antenna assembly of the type described in Sec. 5-6 provides more than sufficient directivity throughout its operating band. Since these antennas are also balanced, they are quite useful as test antennas in pattern measurements.

An unbalanced, *i.e.*, unsymmetrical, antenna is unsatisfactory for several reasons. (1) The direction in which its lobe points will change with frequency, thus requiring constant adjustment of the antenna orientation. (2) The outside of the cable used to feed it is likely (at least at some frequencies) to become excited and radiate. Under this circumstance the direction from which the signal arrives at the antenna being tested, and the polarization of that signal, will depend upon the position of the feed cable and will change whenever any change occurs in the position of the feed cable. This effect is analogous to that produced by the use of a poorly shielded signal source. (3) A difficulty that is more likely to be encountered with unbalanced than with balanced antennas is that of impure polarization. With balanced antennas, the field (at least in the major lobe) is likely to be essentially plane polarized. With unbalanced antennas, this is frequently not the case, and, moreover, the polarization of the field may be expected to vary with frequency. It then becomes impracticable to determine which component of the field of the antenna under test is being measured, and interpretation of the data may be impossible.

The term *balanced antenna* is usually applied to symmetrical dipoles or dipole-reflector assemblies. However, as far as the discussion in the preceding paragraph is concerned, a symmetrical horn is also balanced. In fact a symmetrical horn excited with a $TE_{1,0}$ wave is a particularly convenient test antenna.

2-27. Site and Mechanical Facilities.—The site requirement insofar as freedom from obstruction is concerned has been mentioned above. As to other requirements, the drawing of generalizations is not feasible because of the wide variety of circumstances that may be encountered. Similarly, the mechanical facilities that are most appropriate will depend upon the particular situation. Obviously, remotely controlled turning

mechanisms will usually speed the work. There is, however, one specific practice that is frequently applicable and that has proved very worthwhile. This consists in the use of towers on which to mount both the antenna under test and the transmitting antenna. As with the recommendations made above, the advantage of this is not confined to broadband work, but is merely greater in such work. The object in using towers is to permit directing the radiation so that reflections from the ground (or ground screen) will not reach the receiving antenna. This is not a necessary condition for pattern measurement, but it eliminates the necessity (which otherwise exists) for changing the relative position of the two antennas with each change in frequency. Figure 2-20 shows a setup in which towers are used in this way. Figure 2-21 shows the radiation pattern of a typical antenna that may be used for transmitting. With this pattern the field intensity of the horizontally directed radiation is eight times as

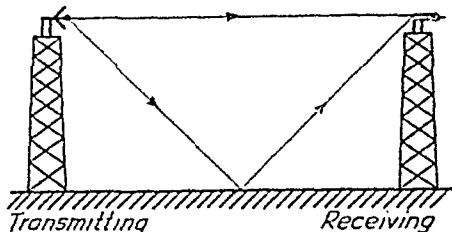


FIG. 2-20.—Antenna towers.

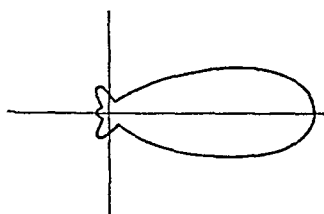


FIG. 2-21.—Radiation pattern of typical antenna used for transmitting.

great as that of the radiation directed 45 deg below the horizontal, which, with the towers shown, is the direction of propagation for the energy that reaches the receiving antenna by reflection. Thus, at the receiving antenna the directly received energy (even neglecting the difference in path lengths) is much greater than the reflected energy. Consequently, the resultant field will for all practical purposes be equal to that of the directly transmitted radiation regardless of the relative phase of the energy arriving along the two paths. If the transmitting lobe were directed slightly above the horizontal, the relative intensity of the reflected energy would be even more reduced. In addition, it is frequently true that the pattern of the antenna under test is such as further to discriminate against the reflected signal. Thus the field at the receiving antenna (on the assumption that radiation from the associated equipment is small) is substantially that which arrives from the transmitting antenna alone. This is a necessary condition for pattern measurement, regardless of frequency. The desirability of towers becomes apparent from a consideration of the behavior of a setup such as that of Fig. 2-22, in which no towers are used. It is assumed that the



FIG. 2-22.—Transmitting and receiving antennas with low or no towers.

ground has approximately uniform reflecting properties (obtained through the use of a covering of wire mesh, if necessary); if this is not the case, accurate pattern measuring is impracticable. The angle between the direct path and the path by reflection is now 5 deg. From Fig. 2-21, it is clear that no orientation of the transmitting antenna that allows a strong direct signal can greatly reduce the relative amplitude of the signal that arrives by reflection. Accordingly, at the receiving antenna the intensity of the field from the transmitting antenna will depend upon the relative phase of the two components. If they are out of phase, so that the resultant is very small, stray radiation from the signal source, cables, small obstructions, etc. (and some such stray radiation is always present) will be relatively large and the data are likely to be meaningless. Of course, it is possible to avoid this condition by appropriately adjusting the spacing and the heights of the antennas. The relations involved are well known,¹ and when understood the adjustments can be made experimentally, rather rapidly, but nevertheless the fact that a separate adjustment is required for each frequency substantially impedes work with broad-band antennas. Accordingly, towers should be used when practicable.

2-28. Automatic Pattern-measuring Methods.—In Sec. 2-22, it was suggested that auxiliary devices that speed the taking of data are an important part of the equipment for measuring radiation patterns. Such devices may take many forms, varying in elaborateness from simple remote controls for the turning mechanisms, to equipment that automatically produces permanent records of the radiation patterns. A relatively simple and very useful system is one in which a polar plot of the radiation pattern is presented upon the screen of a cathode-ray oscilloscope. A system of this type is identical with one of the forms of direction finders described in Sec. 9-13. For pattern-taking purposes, it is merely necessary that the amplitude of the radial deflection be made proportional to the strength of the received signal. This has been satisfactorily achieved by the use of a superheterodyne receiver and the indicating circuit of Sec. 11-17. This equipment, although it does not yield so accurate information in regard to minor lobes as some other systems, is very well suited to many development problems, since it permits one to see instantly the effect of any change made. Permanent records are readily obtained by photographing the screen. The patterns of Chap 4 on Cones and Cylinders were obtained in this way.

¹ See, for example, Terman, F. E., "Radio Engineers' Handbook," pp. 691-693, McGraw-Hill Book Company, Inc., New York, 1943.

CHAPTER 3

IMPEDANCE MATCHING, TRANSFORMERS AND BALUNS

BY J. A. NELSON AND G. STAVIS

3-1. Broad-band Impedance Compensation.—The basic principles of impedance compensation are discussed in Secs. 1-4 and 1-5. Briefly, impedance compensation is accomplished by introducing into the transmission line a reactive network so chosen and so placed that the reflected wave from the reactive network tends to cancel the reflected wave from the antenna. When an antenna impedance is to be compensated at a single frequency, or over a narrow band of frequencies, the reactive network may be designed to give complete compensation. When an antenna impedance is to be compensated over a broad band of frequencies, complete compensation throughout the entire band is usually not feasible; however, the compensating reactive network may be designed so that at all frequencies within the band the reflection coefficient of the compensated antenna is less than a specified limit (see Sec. 1-5). This is consistent with the definition of antenna band width given in Sec. 1-1 as the frequency band within which the reflection coefficient of the antenna does not exceed a certain limit, the value of which depends upon the purpose of the antenna.

The input impedance of an antenna at a given frequency may be expressed as a complex number $R + jX$ or $Z\angle\theta$ and may be plotted as a point on a complex plane, called an R - X diagram, having abscissa R and ordinate X (see Fig. 3-1). Since only positive values of resistance occur, the R - X diagram consists of the first and fourth quadrants of a complete complex plane. The variation of input impedance of an antenna with frequency may be shown graphically on an R - X diagram by plotting the impedance point determined for each frequency and connecting the points with a curve (see Fig. 3-2).

It will be shown in the first part of this chapter that an analysis of the transformation equations of line transformers, parallel networks, and series networks in terms of variable frequency and impedance values

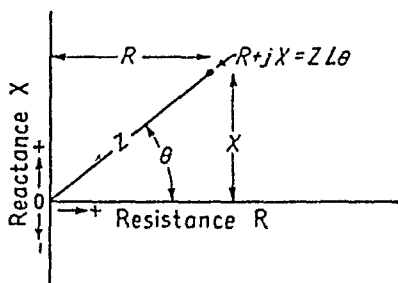


FIG. 3-1.—Location of a point $R + jX$ on an R - X diagram.

leads to mathematical expressions for antenna impedance and compensated input impedance that may be represented as areas on an R - X diagram. Since one may consider a curve of impedance vs. frequency in terms of the area within which it lies, the graphical presentation of the equations in the form of areas is directly and immediately applicable to the problem of impedance compensation.

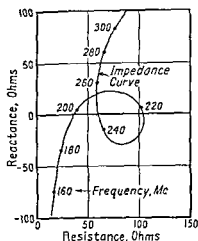


FIG. 3-2 —An impedance curve plotted on an R - X diagram

centered on the R axis, with R intercepts at Z_0/σ and $Z_0 \times \sigma$, where Z_0 is the characteristic impedance of the transmission line. The standing-wave ratio σ is related to the reflection coefficient by the formulas

$$\sigma = \frac{1 + \rho}{1 - \rho} \quad (3-1)$$

$$\rho = \frac{\sigma - 1}{\sigma + 1} \quad (3-2)$$

Figure 3-3 shows several constant ρ circles for $Z_0 = 50$. Within a circle of constant ρ lie all the impedance points that have a lower reflection coefficient. Thus, on an R - X diagram, a constant ρ circle may be drawn, which, since it includes all the acceptable input impedance values for a satisfactory broad-band antenna, we shall call a *definition circle*. The position and radius of the circle are determined by the characteristic impedance of the transmission line and by the limiting value of the reflection coefficient.

Impedance compensation consists of the transformation of an antenna impedance that has a reflection coefficient greater than the specified limit to an input impedance value that has a reflection coefficient less than the specified limit. This process is represented on an R - X diagram

3-2. The Definition Circle.—The input impedance of a broad-band antenna is acceptable when the reflection coefficient is less than the specified limiting value. It can be shown that lines of constant reflection coefficient ρ and hence of constant standing-wave ratio σ are circles on an R - X diagram,

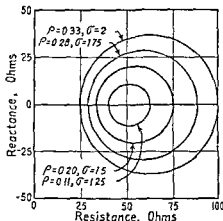


FIG. 3-3 —Several circles of constant reflection coefficient on a 50-ohm line plotted on an R - X diagram.

by the transformation of a point outside the definition circle to a point inside the definition circle and, for a broad-band antenna, by the transformation of the portion of the antenna impedance curve corresponding to the frequency band of interest from a position outside the definition circle to a position within the definition circle.

3-3. Transformation Circles.—The present approach to the problem of impedance compensation may be considered a reversal of the direct approach. The area corresponding to the acceptable input impedance (the definition circle) is stated first. Then the transformation equations are so expressed that areas on the R - X diagram are obtained as functions of the acceptable impedance at the input to the compensating network, with the quantities that describe the compensating network as parameters. General rules are obtained for determining the values of parameters that will transform the definition circle to the area on an R - X diagram that approximates the area within which the antenna impedance curve lies. The values of the parameters correspond to the design constants of the compensating network.

The final transformation equations for the three types of networks considered are all in the standard form of a circle

$$(R - C_1)^2 + (X - C_2)^2 = C_3^2 \quad (3-3)$$

We shall call these circles *transformation circles*. The quantities C_1 , C_2 , and C_3 are functions of the definition circle, which is known, and of the parameters of the compensation network. The transformation circles are completely defined for a particular compensating network at a given frequency by an equation in the form of Eq. (3-3). When frequency varies, Eq. (3-3) represents a family of transformation circles, since the values of the network parameter and, hence, C_1 , C_2 , and C_3 are functions of frequency. It will be shown that the envelope of this family consists of two curves, which we shall call *boundary curves*, between which an antenna impedance curve must lie to be compensated by the network the equation of which defines the boundary curves.

For any particular antenna the impedance-compensation problem may be visualized very quickly by plotting on an R - X diagram the antenna-impedance curve and the appropriate definition circle. By drawing the boundary curves corresponding to various possible networks, it may be determined by inspection (1) whether the impedance curve does not fall completely within any pair of boundary curves, and therefore the antenna impedance cannot be compensated by a single network, or (2) whether the impedance curve does fall within a pair of boundary curves, and therefore it may be possible to compensate the antenna impedance with the network corresponding to the boundary curves. In the second case, it is often possible to determine the exact network

values by drawing one or two of the specific transformation circles, as discussed in Secs. 3-6, 3-8, and 3-9

A specific example of an impedance transformation (shown in Fig. 3-4) is given here merely to illustrate the type of problems with which we are concerned. In the example, the characteristic impedance of the feed line is 50 ohms and the maximum acceptable coefficient of reflection is 0.33, corresponding to the standing-wave ratio of 2:1. The definition circle thus passes through the points $R = 50/2$ and $R = 50 \times 2$. Figure 3-4b shows the impedance presented to the transmission line when

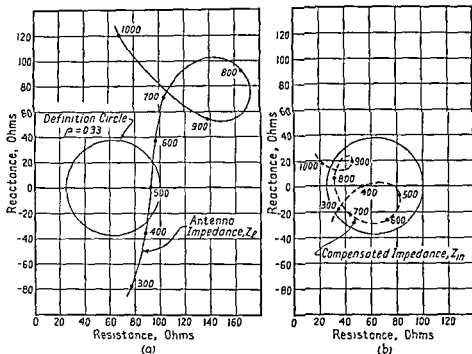


FIG. 3-4—An example of an antenna impedance curve at (a), and at (b) the input impedance curve to which it may be transformed by the appropriate compensating network.

the proper compensating network is used. The network employed, and the techniques involved in designing it, are described in Sec. 3-10.

3-4. Types of Compensating Networks.—There are various possible compensation networks. At lower frequencies it is possible to use lumped-constant compensation networks, such as coils and capacitors, but at frequencies above 100 Mc it is usually more convenient to use distributed-constant compensating networks. Those most frequently used, because of their simplicity, are of three types: (1) the line transformer, (2) the series network, and (3) the parallel network.

A line transformer consists of a short length of transmission line of characteristic impedance different from that of the main transmission

line interposed between the antenna and the transmission line. A series network consists of a short length of open-circuited or short-circuited transmission line placed in series with the inner or outer conductor of the main transmission line. A parallel network consists of a short length of open-circuited or short-circuited transmission line placed in parallel with the antenna.

Several possible methods of construction of these types of compensating network are shown in Fig. 3-5. The transmission lines used for

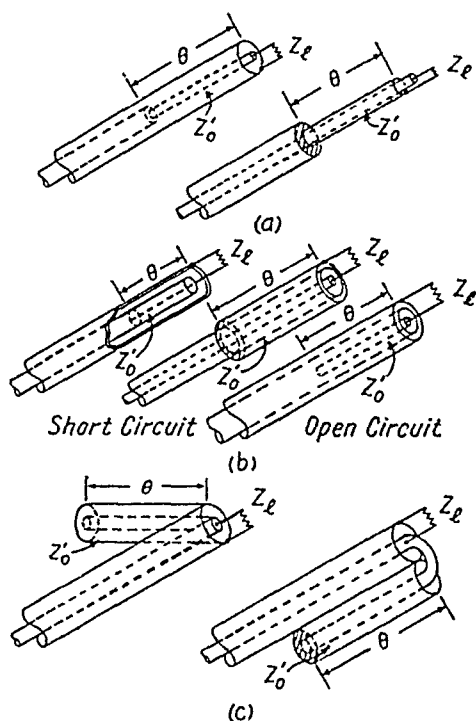


FIG. 3-5.—Examples of the several types of compensating networks in the form in which they can be constructed for use with a coaxial transmission line.

compensating networks are usually short in wavelengths and built of such materials that any loss introduced by the compensating network may be neglected. In all three of the compensating networks considered, there are two variables, the characteristic impedance Z_0' and the electrical length θ . The characteristic impedance Z_0' is a constant for any particular network, but the electrical length θ is a function of frequency. It is usually convenient to consider the effect of a compensating network as the frequency changes in terms of varying θ . Since θ is proportional to frequency for a fixed physical length of transmission line, one may quickly convert θ values to the corresponding frequencies once the relationship at any one value of θ is determined.

3-5. The Line Transformer: Preliminary Considerations.—Since a line transformer is simply a short length of transmission line inserted in series with the antenna, the impedance at the input to the line transformer is related to the antenna impedance by the transmission-line equation

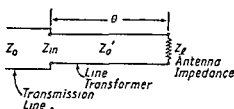


FIG. 3-6.—Diagrammatic representation of a line-transformer compensating network.

$$Z_{in} = Z_0' \frac{Z_L + jZ_0' \tan \theta}{Z_0' + jZ_L \tan \theta} \quad (3-4)$$

where Z_{in} is the impedance at the input to the line transformer, Z_L the impedance of the antenna, Z_0' the characteristic impedance of the line transformer, and θ the electrical length of the line transformer, as illustrated in Fig. 3-6. When Z_L and Z_0' are held constant and θ is varied from 0 to 180 deg, the values of Z_{in} form circles on an R - X diagram, with centers and radii dependent upon the absolute value of the reflection coefficient given by

$$\frac{Z_0' - Z_L}{Z_0' + Z_L}$$

These circles are of course the same as the constant-standing-wave-ratio

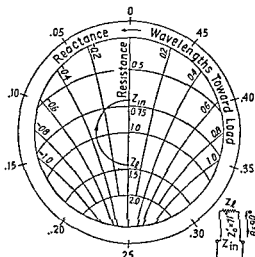


FIG. 3-7.—The transformation of the point Z_L to the point Z_{in} by a quarter-wave line transformer plotted on the central portion of a simplified Smith chart.

circles shown in Fig. 3-3. At the frequency at which $\theta = 90$ deg, the quarter-wave transformer results, and

$$Z_{in} = \frac{(Z_0')^2}{Z_L} \quad (3-5)$$

It is obvious from this relationship that at a single frequency a purely resistive antenna may be matched to a transmission line of real characteristic impedance by means of a quarter-wave transformer. The characteristic impedance Z_0' of the quarter-wave line transformer is the geometric mean between the antenna resistance R_l and the feed-line impedance Z_0 .

$$Z_0' = \sqrt{Z_0 R_l} \quad (3-6)$$

Example: A quarter-wave transformer is illustrated on a Smith chart (see Sec. 2-2) in Fig. 3-7, where R_l is 100 ohms and the transmission line impedance Z_0 is 50 ohms. The line-transformer characteristic impedance Z_0' is, therefore, $\sqrt{50 \times 100}$, or 71, ohms. The ratio of the antenna impedance to the characteristic impedance of the line transformer is plotted on the R -axis at 1.41. This impedance is rotated through the angle corresponding to 0.25λ

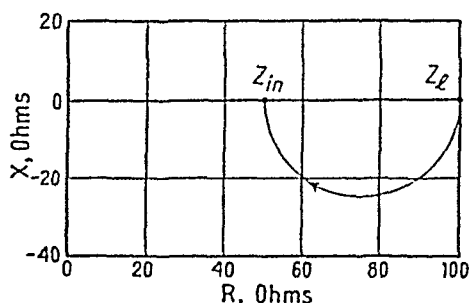


FIG. 3-8.—The transformation of the point Z_l to the point Z_{in} by a quarter-wave line transformer plotted on an R - X diagram.

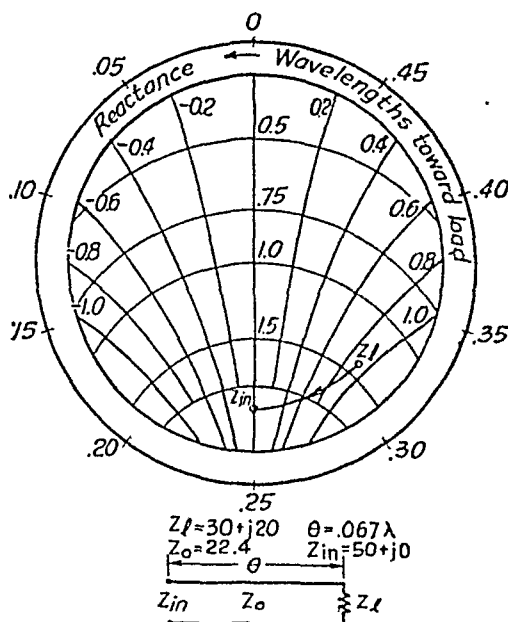


FIG. 3-9.—A transformation on a simplified Smith chart from an antenna impedance Z_l of $30 + j20$ to an input impedance Z_{in} of $50 + j0$.

toward the generator along a circle of constant standing-wave ratio to Z_{in}/Z_0' , at the point 0.71, or $5/7$, corresponding to a Z_{in} of 50 ohms. Figure 3-8 illustrates the same transformation on an R - X diagram. Successive points along the line of rotation give the impedance looking toward the antenna at corresponding

positions along the line transformer, or at corresponding frequencies lower than that at which $\theta = 90$ deg if always $R_i = 100$ ohms.

Many complex load impedances may be matched to the feed line (or to another complex impedance) by a line transformer at a single frequency. The following equations for the line-transformer characteristic impedance Z_0' and the electrical length θ required to transform Z_i to Z_{in} are derived from the transmission-line equation, Eq. (3-4):

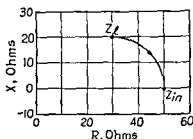


FIG. 3-10—A transformation on an R - X diagram from an antenna impedance Z_i of $30 + j20$ to an input impedance Z_{in} of $50 + j0$.

$$\left. \begin{aligned} Z_0' &= \sqrt{A + \frac{BD}{C}} \\ \tan \theta &= \frac{CZ_0'}{B} \end{aligned} \right\} \quad (3-7)$$

where

$$\begin{aligned} A &= R_i R_{in} - X_i X_{in} \\ B &= R_i X_{in} + R_{in} X_i \\ C &= R_{in} - R_i \\ D &= X_{in} - X_i \end{aligned}$$

Example: A transformation from an antenna impedance of $30 + j20$ to an input impedance of $50 + j0$ is illustrated on a Smith chart in Fig. 3-9 and on an R - X diagram in Fig. 3-10. Upon substituting these values of Z_i and Z_{in} in Eq. (3-7), one finds that $Z_0' = 22.4$ and $\theta = 24$ deg, or 0.067λ .

3-6. The Line Transformer: Transformation Circles.—Let us now consider how a line transformer effects the transformation of circles on an R - X diagram to the definition circle. First we shall disregard the effect of changing frequency and determine the equation for those circles which may be transformed to the definition circle by a line transformer with a given electrical length θ . Let us allow Z_{in} to assume all the values within the definition circle, or for the sake of simplicity, all the values on the definition circle. The equation for the definition circle is

$$(R_{in} - a)^2 + X_{in}^2 = b^2 \quad (3-8)$$

where the point $R = a$, $X = 0$ is the center and b is the radius of the circle, as shown on Fig. 3-11. The quantities a and b are related to the standing-wave ratio σ and the line characteristic impedance Z_0 by the equations

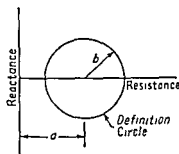


FIG. 3-11.—Construction of definition circle, where

$$a = \frac{Z_0 \sigma + Z_0 / \sigma}{2}$$

and

$$b = \frac{Z_0 \sigma - Z_0 / \sigma}{2}$$

$$\begin{aligned} a &= \frac{Z_0}{2} \left(\sigma + \frac{1}{\sigma} \right) \\ b &= \frac{Z_0}{2} \left(\sigma - \frac{1}{\sigma} \right) \end{aligned} \quad (3-9)$$

Equation (3-8) may be rewritten

$$Z_{in} \bar{Z}_{in} - a(Z_{in} + \bar{Z}_{in}) = b^2 - a^2 \quad (3-10)$$

where

$$Z_{in} = R_{in} + jX_{in}$$

and

$$\bar{Z}_{in} = R_{in} - jX_{in}$$

The values of Z_{in} and \bar{Z}_{in} in terms of Z_l and \bar{Z}_l may be obtained from Eq. (3-4) and are

$$\begin{aligned} Z_{in} &= Z_0' \frac{Z_l + jZ_0' \tan \theta}{Z_0' + jZ_l \tan \theta} \\ \bar{Z}_{in} &= Z_0' \frac{\bar{Z}_l - jZ_0' \tan \theta}{Z_0' - j\bar{Z}_l \tan \theta} \end{aligned} \quad (3-11)$$

Substituting in Eq. (3-10) the values of Z_{in} and \bar{Z}_{in} given in Eq. (3-11) we have

$$\begin{aligned} &\left(Z_0' \frac{Z_l + jZ_0' \tan \theta}{Z_0' + jZ_l \tan \theta} \right) \left(Z_0' \frac{\bar{Z}_l - jZ_0' \tan \theta}{Z_0' - j\bar{Z}_l \tan \theta} \right) \\ &- a \left(Z_0' \frac{Z_l + jZ_0' \tan \theta}{Z_0' + jZ_l \tan \theta} + Z_0' \frac{\bar{Z}_l - jZ_0' \tan \theta}{Z_0' - j\bar{Z}_l \tan \theta} \right) = b^2 - a^2 \end{aligned} \quad (3-12)$$

Equation (3-12) may be changed to the form

$$\begin{aligned} &\left[R_l - \frac{a(1 + \tan^2 \theta)}{1 + \frac{a^2 - b^2}{(Z_0')^2} \tan^2 \theta} \right]^2 + \left[X_l - \frac{Z_0' \tan \theta \left(\frac{a^2 - b^2}{(Z_0')^2} - 1 \right)}{1 + \frac{a^2 - b^2}{(Z_0')^2} \tan^2 \theta} \right]^2 \\ &= \left[\frac{b(1 + \tan^2 \theta)}{1 + \frac{a^2 - b^2}{(Z_0')^2} \tan^2 \theta} \right]^2 \end{aligned} \quad (3-13)$$

which is the standard form of the equation of a circle, since the quantities a , b , and Z_0' are all constant for a particular line transformer and definition circle and θ is constant at a given frequency. The center of the transformed circle is at

$$R = \frac{a(1 + \tan^2 \theta)}{1 + \frac{a^2 - b^2}{(Z_0')^2} \tan^2 \theta} \quad X = \frac{Z_0' \tan \theta \left(\frac{a^2 - b^2}{(Z_0')^2} - 1 \right)}{1 + \frac{a^2 - b^2}{(Z_0')^2} \tan^2 \theta}$$

and the radius is

$$\frac{b(1 + \tan^2 \theta)}{1 + \frac{a^2 - b^2}{(Z_0')^2} \tan^2 \theta}$$

The circle of Eq. (3-13), which is a transformation circle, is transformed to the definition circle by the transformation equation of a line transformer of characteristic impedance Z_0' and length θ .

Example: Figure 3-12 is an example in which the definition circle is that for a 1.1:1 standing-wave ratio on a 100-ohm line, and passes through the points

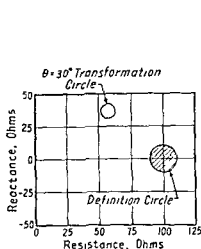


FIG. 3-12.—A transformation circle and a definition circle plotted on an R - X diagram, where the definition circle is a 1.1:1 standing-wave-ratio circle on a 100-ohm line, $Z_0' = 50$ ohms and $\theta = 30$ deg

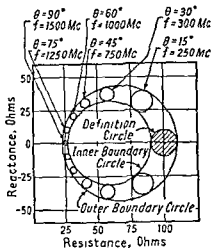


FIG. 3-13.—A family of transformation circles for a line transformer of 50-ohm characteristic impedance, where the definition circle is a 1.1:1 standing-wave-ratio circle on a 100-ohm line.

$R = 110$, $X = 0$, and $R = 91$, $X = 0$. The quantities a and b thus have the values 100.5 and 9.5. In this example, the line transformer is 30 deg long and has a characteristic impedance Z_0' of 50 ohms. Substituting the given values of a , b , Z_0' , and θ in Eq. (3-13), we find that the transformation circle is centered at $R = 57.5$, $X = 37$ and has a radius of 5.45. The impedance looking into the line transformer described above lies within the definition circle when the antenna impedance at the end of the line transformer lies within the transformation circle.

A line transformer has a given length θ at one frequency only, but a , b , and Z_0' are independent of frequency. Thus, a family of transformation circles with θ as the parameter represents those impedances that will be transformed to impedances lying within the definition circle by a given line transformer as the frequency is varied (see Fig. 3-13). Thus, when a line transformer is 30 deg long at 500 Mc, it is 90 deg long at 1500 Mc, etc.

Figure 3-14 is a family of transformation circles for a line transformer of 100-ohm characteristic impedance, where the definition circle is a 2:1 standing-wave-ratio circle on a 50-ohm transmission line.

Several general relationships may be found in Eq. (3-13) and are clearly illustrated in Figs. 3-13 and 3-14.

1. The transformation circles are not all the same size, the circles becoming larger as they move into the higher resistance area of the diagram. In fact, the radius of each circle is always directly proportional to the resistance value of the center of that circle, as is shown by Eq. (3-13).

2. Each series of transformation circles is bounded by two circles which, as mentioned in Sec. 3-3, are called boundary curves. One is tangent to the resistance-maximum point of the definition circle and the resistance-minimum point of the $\theta = 90$ -deg circle, the other is tangent to the resistance-minimum point of the definition circle and the resistance-maximum point of the $\theta = 90$ -deg circle. In Fig. 3-14, the inner boundary circle has degenerated to a point.

When the definition circle and the $\theta = 90$ -deg circle overlap, the inner boundary circle is contained within all the transformation circles, and consequently any point lying within an *inverted inner boundary circle* remains within it irrespective of the value assigned to θ . Figure 3-15 shows such a case.

3. The transformation circles progress within the boundary curves in a counter-clockwise direction as θ increases.

3-7. Impedance-curve Transformation by Line-transformer Methods. The properties of the transformation circles suggest a simple technique for establishing the suitability of any given impedance curve for compensation by a line transformer. When any impedance curve and any definition circle are plotted on an R - X dia-

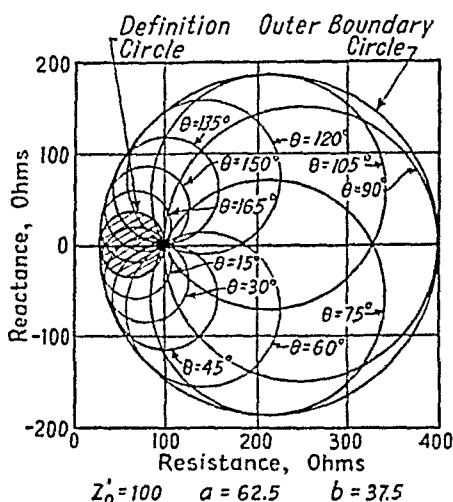


FIG. 3-14.—A family of transformation circles for a line transformer of 100-ohm characteristic impedance, where the definition circle is a 2:1 standing-wave-ratio circle on a 50-ohm line.

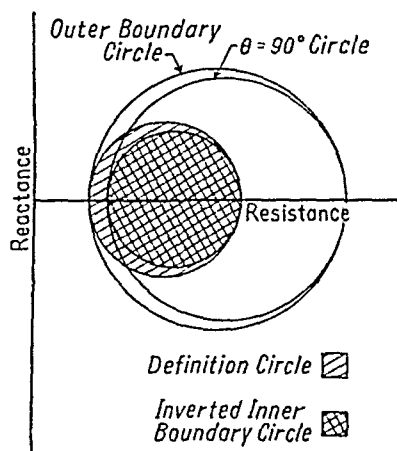


FIG. 3-15.—An example of an inverted inner boundary circle occurring when the characteristic impedance of the line transformer lies within the definition circle.

impedance curve for compensation by a line transformer. When any impedance curve and any definition circle are plotted on an R - X dia-

gram, an appropriate inner or outer boundary circle may be constructed by inspection. The other boundary circle may then be drawn, the ratio between the pure-resistance intercepts of the boundary circles being kept the same for the $\theta = 90$ -deg position as for the definition-circle position.

Example: In the construction of Fig. 3-16 a definition circle and the impedance curve are drawn first. On the assumption that it is desirable to compensate the curve between 100 and 190 Mc, an outer boundary circle may then be drawn, including the 100- and 190-Mc impedance points, which is tangent to the definition circle at the low-resistance intercept. The portions of the impedance curve lying outside this boundary cannot be compensated. The inner boundary circle may now be drawn. The definition circle crosses the R -axis at $R = 22$ and $R = 53$, thus the ratio of intercepts is 2.4 ($Z_0 = 34$ ohms, $\sigma = 1.55$). The outer boundary circle crosses the R -axis at 170. Since the ratio of intercepts for the $\theta = 90$ -deg circle is the same as for the definition circle, the inner boundary circle will cross the R -axis at $170/2.4$ or at $R = 71$. It may be possible to compensate with a line transformer for that part of the impedance curve which falls completely within the area bounded by the two boundary circles. It is clear from Fig. 3-16 that the impedance curve does fall within the area between the boundary circles between 100 and 190 Mc and, hence, may be susceptible to compensation by a line transformer consisting of a transmission line of characteristic impedance equal to the mean

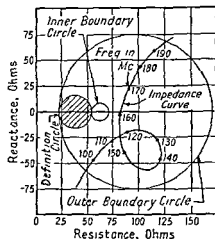


FIG. 3-16.—The construction of appropriate boundary circles for a given impedance curve and definition circle.

between the two resistance intercepts of either of the boundary circles. Thus $\sqrt{170 \times 22}$ ohms and $\sqrt{71 \times 53}$ ohms are both equal to 61 ohms, the characteristic impedance of the required line transformer.

To determine what length of transmission line should be used and what portion of the impedance curve lying within the boundaries will be compensated by the line transformer requires that the effect of the varying parameter θ be investigated. A line transformer has a different electrical length θ for every frequency. Therefore, a given line transformer will compensate for one transformation circle at one frequency and another transformation circle at another frequency.

Example: A line transformer is 30 cm long and of 100-ohm characteristic impedance (see Fig. 3-17). At 250 Mc, 30 cm is 90 deg; so at 250 Mc the line transformer will compensate the impedances lying within the $\theta = 90$ -deg circle.

At 125 Mc the line transformer is 45 deg long and will therefore compensate the impedance lying within the $\theta = 45$ -deg circle.

Each point on an impedance curve corresponds to one frequency. If the impedance point for a frequency falls within the θ circle for that frequency, the point may be compensated. When, at each successive frequency, the impedance point falls within the corresponding θ circle, the whole curve may be compensated. The impedance curve may be so shaped that it tracks with the θ circles, or they may be in opposition (see Fig. 3-18). In Fig. 3-18a, most of the curve within the boundary circles may be compensated, since the increasing-frequency points fall within circles of increasing θ ; in Fig.

3-18b only a small part of the curve within the boundary circles may be compensated, since increasing-frequency points fall in decreasing- θ circles.

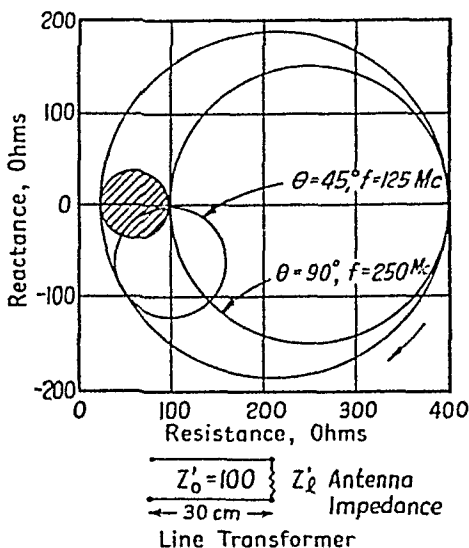


FIG. 3-17.—Transformation circles in which a given physical length is assigned to the line transformer, making the circles of constant θ also circles of constant frequency.

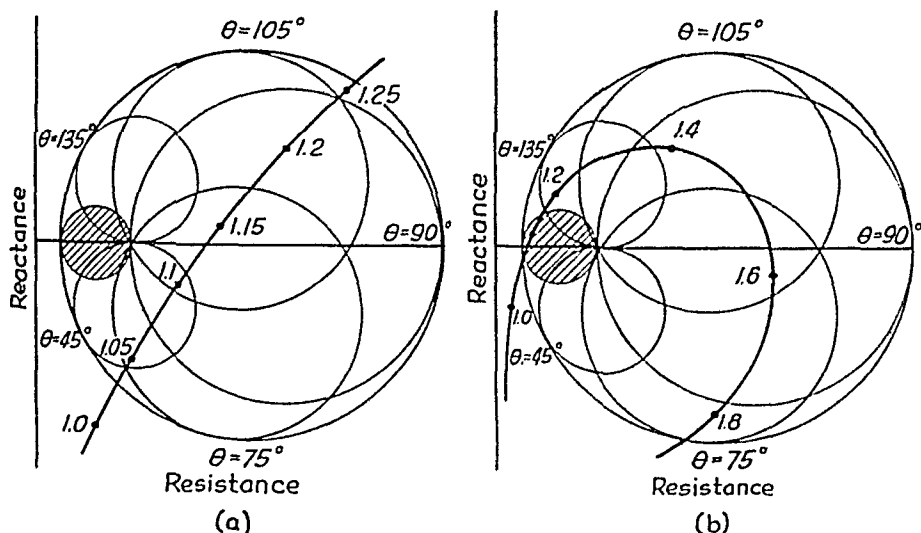


FIG. 3-18.—Two impedance curves drawn with boundary lines and transformation circles. The curve in a tracks with the transformation circles; that in b is in opposition.

Now let us go back to the impedance curve of Fig. 3-16. The next step in determining the appropriate values for the line transformer is to

draw in the various θ circles. It is possible to locate the $\theta = 45$ -deg and $\theta = 135$ -deg circles in a very simple manner. A circle, centered at the origin, with a radius equal to Z_0' intercepts the two boundary circles at the points where the $\theta = 45$ -deg and $\theta = 135$ -deg transformation circles

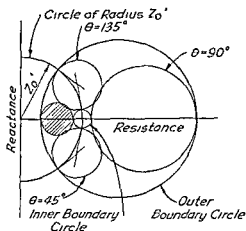


Fig. 3-19 — Construction of $\theta = 45$ -deg and $\theta = 135$ -deg transformation circles.

are tangent to the boundary circles (Fig. 3-19). The centers of the $\theta = 45$ -deg and $\theta = 135$ -deg circles may be located by constructing the perpendiculars to the tangents of the boundary circles at the points where they are crossed by the circle of radius Z_0' . The $\theta = 90$ -deg circle is, of course, centered on the R -axis and tangent to both the inner and outer boundary circles, and the $\theta = 0$ -deg and $\theta = 180$ -deg circles coincide with and are identical to the definition circle.

Knowing the position of four of the transformation circles, $\theta = 0$ deg, $\theta = 45$ deg, $\theta = 90$ deg, and $\theta = 135$ deg, it is usually possible to interpolate with fair accuracy the values of θ required to compensate other positions within the region between the boundary circles.

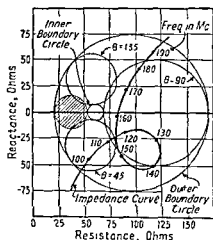


Fig. 3-20 — Impedance curve of Fig. 3-16 plotted with transformation circles.

Figure 3-20 is the impedance curve of Fig. 3-16 plotted on an R - X plane with the definition circle shaded, the $\theta = 45$ -deg, $\theta = 90$ -deg, and $\theta = 135$ -deg transformation circles and the inner and outer boundary circles drawn and labeled. Before assigning a value to θ for a particular frequency, one should consider the critical portions of the impedance

curve. Points lying on one of the boundary circles can be compensated only by one value of θ , since there is only one θ -circle tangent to the boundary circle at a given point. Midway between the two boundary circles, the θ -circles overlap, and a single point may be compensated by many values of θ . We see, therefore, that the portions of an impedance curve which lie on or near the boundary circle may be considered critical. Any portion of the impedance curve that moves in a clockwise direction, *i.e.*, where, as the frequency increases, the θ -circles decrease, is also critical. It is usually wise to assign a value to θ for a frequency at which the impedance curve is critical and from that value to determine the values of θ for the frequencies of other points on the curve. There are three critical points to be considered on the impedance curve of Fig. 3-20, the 100- and 190-Mc points and the 140-Mc point. The 100- and 190-Mc points are critical because they lie near the outer boundary circle. The 140-Mc point is critical for two reasons: (1) because it is quite close to the outer boundary circle and (2) because between about 125 and 150 Mc the impedance curve moves in a clockwise direction. Choice of the optimum θ -value for the 140-Mc point gives the best chance of compensating for both the 125-Mc point and the 150-Mc point. The 140-Mc point is probably closest to the center of the $\theta = 75$ -deg circle. A line transformer 75 deg long at 140 Mc has the lengths shown in Table 3-1 at the other frequencies.

TABLE 3-1

Frequency	θ	$\theta/360$
100	54	0.150
110	59	0.164
120	64	0.178
130	70	0.194
140	75	0.208
150	80	0.222
160	86	0.238
170	91	0.253
180	96	0.267
190	102	0.283

It is possible to estimate the position of the θ circles for each value of θ listed in Table 3-1. Each impedance point lying within the corresponding estimated θ circle can be transformed to the definition circle. An inspection of Fig. 3-20 and Table 3-1 shows that this impedance curve can be transformed to the definition circle by a line transformer of 61-ohm characteristic impedance that is 75 deg long at 140 Mc.

If necessary, the exact value of the impedance at the input to the line

transformer may be determined graphically for each frequency on a Smith chart, as shown in Fig. 3-21. For this purpose it is convenient to have printed Smith-chart forms. The antenna-impedance curve rela-

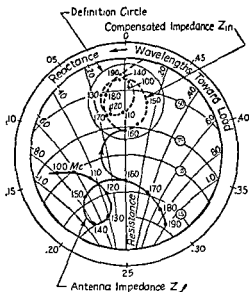


FIG. 3-21.—Transformation of an impedance curve on a Smith chart with the transformation line for the $f = 160$ -Mc point drawn

tive to 61 ohms $\left(\frac{R + jX}{61}\right)$ is plotted on a Smith chart and each point rotated clockwise the appropriate distance in wavelengths (see Table 3-1), as shown for the $f = 160$ -Mc point.

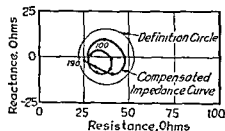


FIG. 3-22.—Transformed impedance curve of Fig. 3-20 plotted on an R - X diagram.

Figure 3-22 is the transformed impedance curve plotted on an R - X diagram.

3-8. The Series Network.—A series network consists of an open-circuited or short-circuited transmission line in series with either the inner or outer conductor of the main transmission line. The input reactance of series networks may be added directly to the antenna impedance, as in the formula

$$Z_i + jX_s = R_i + j(X_i + X_s) \quad (3-14)$$

where $Z_i = R_i + jX_i$ is the antenna impedance and jX_s is the input reactance of the series network.

Since the definition circle equation is

$$(R - a)^2 + X^2 = b^2 \quad (3-8)$$

and a series network adds reactance directly, the equation for the circles which will be transformed to the definition circle by a series network at any one frequency is

$$(R - a)^2 + (X + X_s)^2 = b^2 \quad (3-15)$$

It is clear that this is the equation of a family of circles having the same radius as the definition circle, with the center on the line $R = a$ displaced from the R -axis a distance equal to $-X_s$. A family of such circles is shown in Fig. 3-23.

The reactance jX_s of a series network is a function of the electrical length and characteristic impedance of the line. For a short-circuited line

$$X_s = Z_0' \tan \theta \quad (3-16)$$

and for an open-circuited line

$$X_s = -Z_0' \cot \theta \quad (3-17)$$

where Z_0' is the characteristic impedance and θ is the electrical length of the line.

Since the electrical length of the series network θ is a function of frequency, for a network of given physical dimensions X_s varies with the frequency.

In general the values that X_s has for certain values of θ may be determined by inspection from Eqs. (3-16) and (3-17).

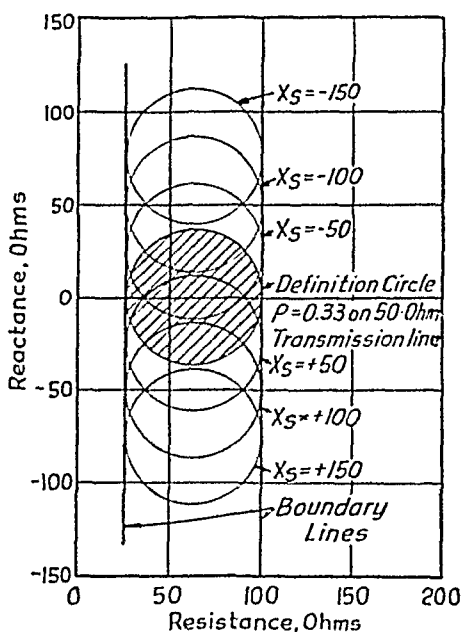


FIG. 3-23.—A family of series-network transformation circles.

1. The value of X_s is infinite for a short-circuited series network 90 deg long and for an open-circuited series network 180 deg long. Consequently, series networks are not useful for impedance compensation at frequencies where θ has such values.

2. The quantity X_s has a value of zero for a half-wave short-circuited series network or a quarter-wave open-circuited series network. At frequencies where θ has these values, the series network does not affect the load impedance. At lower frequencies, the series network adds capacitance, compensating for transformation circles in the inductive region, and at higher frequencies the series network adds inductance, compensating for transformation circles in the capacitive region. Thus, an impedance curve of the shape of Fig. 3-24 may be compensated by an open-circuited series network of length 90 deg at 350 Mc.

3. A short-circuited series network less than 90 deg long or an open-circuited series network between 90 and 180 deg long has inductive reactance. The abso-

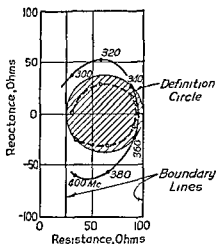


FIG. 3-24.—Transformation of the solid-line impedance curve to the dashed-line impedance curve by an open-circuited 150-ohm line 90 deg long at 350 Mc in series.

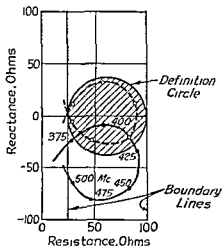


FIG. 3-25.—Transformation of the solid-line impedance curve to the dashed-line impedance curve by a short-circuited 50-ohm line 45 deg long at 475 Mc in series.

lute value of X , increases with θ in these regions, so that the displacement distance of impedance points is a direct function of frequency. Thus a curve such as the one in Fig. 3-25 may be compensated by such a series network.

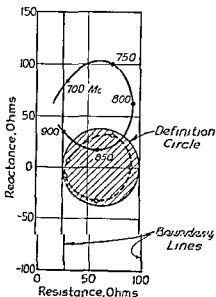


FIG. 3-26.—Transformation of the solid-line impedance curve to the dashed-line impedance curve by an open-circuited 100-ohm line 60 deg long at 800 Mc in series.

with various values of Z_0' or θ . For instance, if we want Z_0' to be 100 ohms, θ

4. An open-circuited series network less than 90 deg long, or a short-circuited series network between 90 and 180 deg long, has capacitive reactance. The absolute value of X , decreases with θ and frequency in these regions, so that the displacement distance of impedance points is an inverse function of frequency. Thus an impedance curve such as the one in Fig. 3-26 may be compensated by such a series network.

5. The rate at which X , changes in value with frequency may be controlled by adjusting the value of Z_0' . This fact is most quickly understood by considering an example. The value of X , that will best compensate for the $f = 300$ -Mc impedance point on the curve in Fig. 3-27a is about $+j40$, since transposing this point 40 ohms straight up places it within the definition circle near the R -axis. This value may be achieved

must be 22 deg, or if θ is to be 50 deg, Z_o' must be 34 ohms. Table 3-2 shows the value of X_s for the two networks at several frequencies, and it is clear that although at 300 Mc the value of X_s is equal for the two networks at other frequencies the values deviate widely.

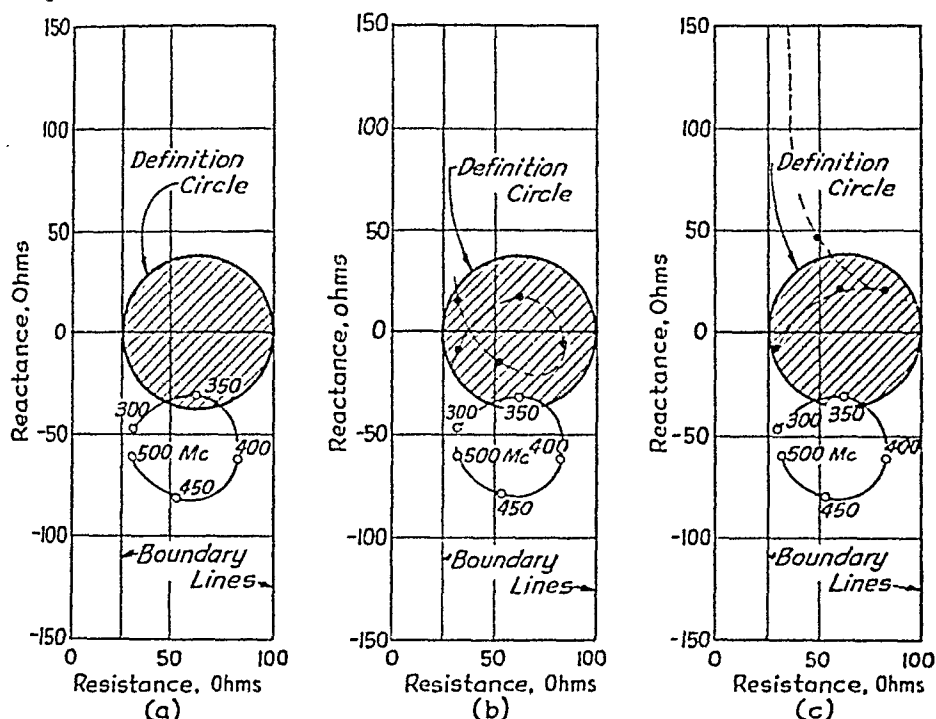


FIG. 3-27.—(a) An arbitrary impedance curve. (b) The effect on the impedance curve of adding in series a short-circuited 100-ohm line 22 deg long at 300 Mc. (c) The effect of adding in series a 34-ohm line 50 deg long at 300 Mc.

The effect of the series networks is shown in Fig. 3-27*b* for $Z_o' = 100$ ohms and in *c* for $Z_o' = 34$ ohms. Obviously, the 100-ohm series network is effective, and the 34-ohm series network is not.

TABLE 3-2

Frequency, Mc	100-ohm series network		34-ohm series network	
	0	$+jZ_o' \tan \theta$	0	$+jZ_o' \tan \theta$
300	22	$j40$	50	$j40$
350	26	$j49$	58	$j55$
400	29	$j56$	67	$j80$
450	33	$j65$	75	$j127$
500	37	$j76$	83	$j280$

A series network is most effective when used to compensate an impedance curve in which the low-frequency points are more inductive, or less

capacitive, than the high-frequency points. A series network tends to close such a curve or to cross the ends of it. An extreme illustration of this effect is the example shown in Fig. 3-27c. A series network of high characteristic impedance distorts the shape of the impedance curve less than a series network of low characteristic impedance, as is clear from a comparison of Figs. 3-27b and c.

3-9. The Parallel Network.—A parallel network consists of a short-circuited or open-circuited transmission line in parallel with the antenna. The electrical length of the parallel network is θ , and its characteristic impedance is Z_0' . The effect of a parallel network of input reactance X_p on an antenna impedance Z_i may be determined by adding the susceptance of the parallel network to the admittance of the antenna, as in the equation

$$Y_i + jB_p = G_i + j(B_i + B_p) \quad (3-18)$$

where Y_i is the admittance of the antenna and B_p is the susceptance of the parallel network. It is obvious from Eq. (3-18) that an admittance may be perfectly matched to the transmission line when $G_i = Y_0$, the admittance of the transmission line, and B_p is chosen equal to $-B_i$.

It is possible to fulfill this condition at a single frequency in any case where the antenna admittance has a conductance component not equal to zero or infinity by placing the parallel network at a point along the transmission line where the conductance $G = Y_0$. The proper position and susceptance of the shunt network may be determined graphically on a Smith chart.

Let us consider an impedance of $80 - j40$ ohms that is to be matched to 50 ohms pure resistance. On the Smith chart, Fig. 3-28, are plotted the Z_i point $(80 - j40)/50$ and 180 deg away from it Y_i , $50/(80 - j40)$. Y_i is rotated toward the generator to the point where $G = 1.0$, a distance the chart shows to be 0.105λ . The addition of a susceptance of $-j0.8 \times 50$ will transform the admittance to the point $G = 1.0$, $B = 0$, and the antenna will be matched to 50 ohms.

Before determining the circles on an R - X diagram within which lie the impedances that may be transformed to impedances plotted within the definition circle by a parallel network, let us consider the transformation on a G - B admittance diagram.

The admittance equation for a parallel network is of the same form as the impedance equation for a series network. Thus the family of admittance circles compensated by a shunt network when plotted on an admittance diagram will be in the same position as the family of impedance circles compensated by a series network when plotted on an impedance diagram (Fig. 3-23). The equation for this family of circles on an admit-

tance diagram is

$$(G - \alpha)^2 + (B - B_p)^2 = \beta^2 \quad (3-19)$$

where α and β are the abscissas of the center and the radius of the admittance definition circle, respectively. This family of circles may be

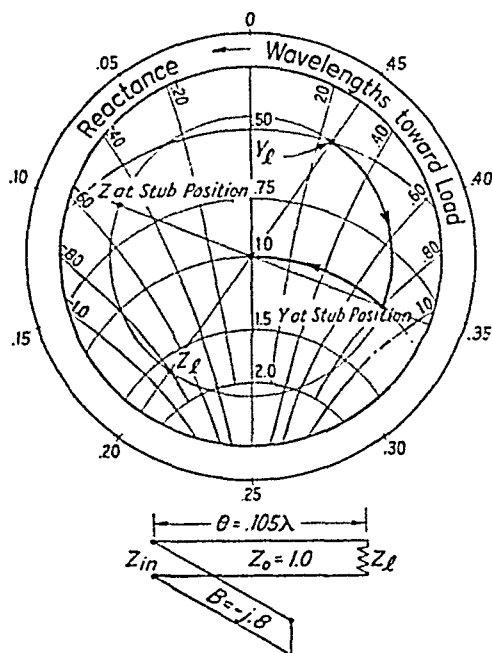


FIG. 3-28.—The use of a parallel network to transform an impedance of $80 - j40$ to an impedance of $50 + j0$ illustrated on a simplified Smith chart.

plotted on an impedance diagram by substituting $R/(R^2 + X^2)$ for G , $-X/(R^2 + X^2)$ for B , a/Z_0^2 for α , and b/Z_0^2 for β , where Z_0 is the characteristic impedance of the main transmission line.¹ Making these sub-

¹ The relationships between a and α , and b and β (the centers and radii of the impedance and admittance definition circles, respectively) may be obtained from the following equations:

$$\left. \begin{aligned} a &= \frac{Z_0}{2} \left(\sigma + \frac{1}{\sigma} \right) & \alpha &= \frac{Y_0}{2} \left(\sigma + \frac{1}{\sigma} \right) \\ b &= \frac{Z_0}{2} \left(\sigma - \frac{1}{\sigma} \right) & \beta &= \frac{Y_0}{2} \left(\sigma - \frac{1}{\sigma} \right) \end{aligned} \right\} \quad (3-9a)$$

Substituting $1/Z_0$ for Y_0 and simplifying transforms the equations to

$$\left. \begin{aligned} \frac{a}{Z_0} &= \frac{\sigma^2 + 1}{2\sigma} & Z_0 \alpha &= \frac{\sigma^2 + 1}{2\sigma} \\ \frac{b}{Z_0} &= \frac{\sigma^2 - 1}{2\sigma} & Z_0 \beta &= \frac{\sigma^2 - 1}{2\sigma} \end{aligned} \right\} \quad (3-9b)$$

Thus $\alpha = \frac{a}{Z_0^2}$ and $\beta = \frac{b}{Z_0^2}$.

stitutions and setting B_p equal to $-1/X_p$ gives the relation

$$\left(\frac{R}{R^2 + X^2} - \frac{a}{Z_0^2}\right)^2 + \left(\frac{X}{R^2 + X^2} + \frac{1}{X_p}\right)^2 = \left(\frac{b}{Z_0^2}\right)^2 \quad (3-20)$$

Equation (3-20) may be rewritten in the form

$$\left(R - \frac{aZ_0^2}{a^2 - b^2 + \frac{Z_0^4}{X_p^2}}\right)^2 + \left(X - \frac{\frac{Z_0^4}{X_p}}{a^2 - b^2 + \frac{Z_0^4}{X_p^2}}\right)^2 = \left(\frac{bZ_0^2}{a^2 - b^2 + \frac{Z_0^4}{X_p^2}}\right)^2 \quad (3-21)$$

Since $a^2 - b^2 = Z_0^2$, this equation may be rewritten¹

$$\left[R - \frac{a}{1 + \left(\frac{Z_0}{X_p}\right)^2}\right]^2 + \left[X - \frac{\frac{Z_0^2}{X_p}}{1 + \left(\frac{Z_0}{X_p}\right)^2}\right]^2 = \left[\frac{b}{1 + \left(\frac{Z_0}{X_p}\right)^2}\right]^2 \quad (3-22)$$

This is the equation of a circle (a transformation circle) since for a given parallel network at a single frequency, a , b , X_p , and Z_0 are all constants. Any impedance point falling within this circle will be transformed to an impedance point that falls within the definition circle by use of a parallel network having an input impedance X_p . The center of the circle is at the point

$$R = \frac{a}{1 + \left(\frac{Z_0}{X_p}\right)^2}, \quad X = \frac{\frac{Z_0^2}{X_p}}{1 + \left(\frac{Z_0}{X_p}\right)^2}$$

and the radius is

$$\frac{b}{1 + \left(\frac{Z_0}{X_p}\right)^2}$$

Figure 3-29 shows a transformation circle that may be transformed to the definition circle by the parallel-network-transformation equation where X_p is assigned a value of -50 . The definition circle is a 2:1-standing-wave-ratio circle on 50-ohm line, passing through the points $R = 100$, $X = 0$ and $R = 25$, $X = 0$.

The quantity a thus has a value of 62.5, and b is equal to 37.5. Substituting $a = 62.5$, $b = 37.5$, $Z_0 = 50$, and $X_p = 50$ in Eq. (3-22), one

¹ Since from Eq. (3-9)

$$a + b = \sigma Z_0 \quad \text{and} \quad a - b = \frac{Z_0}{\sigma}$$

$$a^2 - b^2 = Z_0^2$$

finds that the center of the transformation circle is at the point $R = 31.25$, $X = 25$, and the radius is 19. Figure 3-30 is a family of transformation circles in which the value of X_p is the parameter.

The input reactance X_p of a parallel short-circuited-line network is

$$X_p = Z_0' \tan \theta \quad (3-23)$$

and for an open-circuited line

$$X_p = -Z_0' \cot \theta \quad (3-24)$$

where Z_0' is the characteristic impedance of the parallel network and θ is its electrical length. For a parallel network of given dimensions, θ

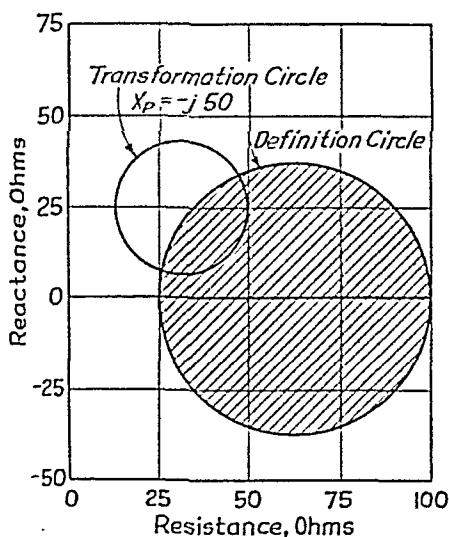


FIG. 3-29.—A transformation circle which may be transformed to the definition circle by a parallel network whose impedance $X_p = -j50$.

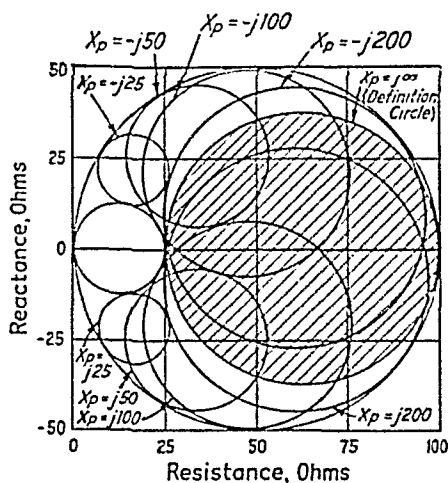


FIG. 3-30.—A family of transformation circles which may be transformed to the definition circle by parallel networks with the indicated X_p values.

varies directly with frequency and Z_0' is constant. At some frequency, therefore, a given parallel network assumes the reactance value corresponding to each transformation circle in Fig. 3-30.

Several general conclusions may be drawn from Eqs. (3-22), (3-23), and (3-24).

1. All the transformation circles are tangent to two *boundary circles* that pass through the origin and are tangent to the definition circle at its R intercepts. Any points lying between these two boundary circles may be transformed to the definition circle, and any points not lying between the boundary circles cannot be transformed to the definition circle. Several examples are shown in Fig. 3-31.

2. Transformation circles the centers of which lie in the capacitive region can be transformed to the definition circle by parallel networks with inductive react-

ance, and transformation circles the centers of which lie in the inductive region can be transformed to the definition circle by parallel networks with capacitive reactance.

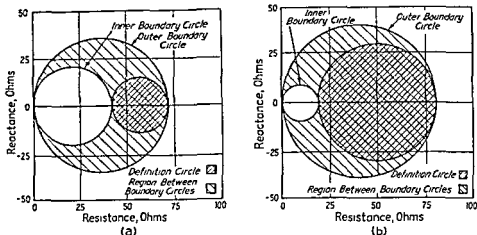


FIG. 3-31.—Examples of the parallel-network inner and outer boundary circles for two different definition circles

3. The ratio between the radius and the resistance value of the center of each transformation circle is constant. Thus the circles become smaller as they progress from the definition circle to the origin.

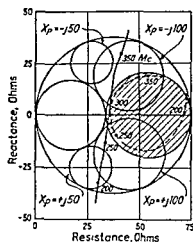


FIG. 3-32.—Transformation of the solid-line impedance curve to the dotted-line impedance curve by an open-circuited parallel network of 100-ohm characteristic impedance 180 deg long at 300 Mc.

4. The value of X_p is zero for a short-circuited network 180 deg long, or an open-circuited network 90 deg long. Consequently, a parallel network is not useful at frequencies where θ has these lengths.

5. The value of X_p is infinite for a short-circuited network 90 deg long, or an open-circuited network 180 deg long, and a parallel network does not affect the antenna impedance at frequencies where θ has these values. In each case, at lower frequencies the value of θ is less and the network has inductive reactance, and at higher frequencies the value of θ is greater and the network has capacitive reactance. Thus, impedance curves such as that in Fig. 3-32 can be compensated by a parallel network.

6. Impedance curves such as those in Figs. 3-33 and 3-34 where all the impedance

points to be compensated lie in either (1) the capacitive region or (2) the inductive region may be compensated, respectively, by (1) short-circuited networks less than 90 deg long or open-circuited networks between 90 and 180 deg

long or (2) short-circuited networks between 90 and 180 deg long or open-circuited networks less than 90 deg long.

7. The rate at which X_p varies in value with frequency may be controlled by adjusting the values of Z_0' and θ . The reactance of a network plotted against frequency is, of course, a tangent-function curve with a constant coefficient equal to the characteristic impedance of the network. The slope of the tangent curve for a given value of θ is thus a function of the characteristic impedance of the network. Figure 3-35 contains two graphs of the reactance of networks vs. frequency where θ has lengths such that the reactance of the network is inductive. The first, in Fig. 3-35a, shows curves of input reactance against frequency for two

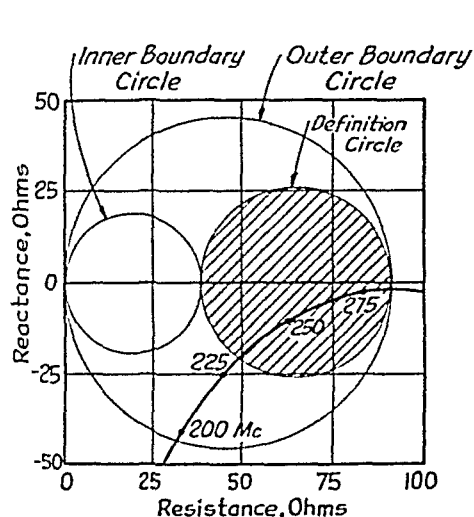


FIG. 3-33.—An impedance curve which requires an inductive parallel network for compensation.

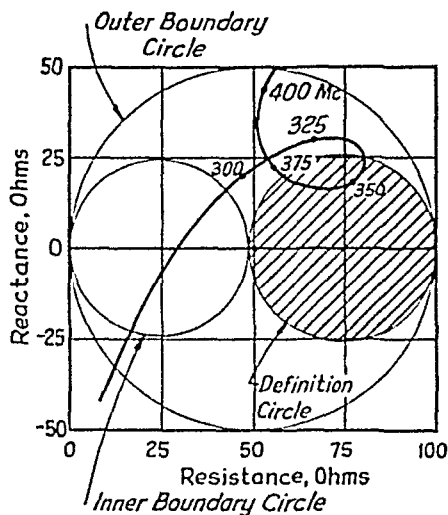


FIG. 3-34.—An impedance curve which requires a capacitive parallel network for compensation.

networks, both of which have a reactance of $+j50$ ohms at 500 Mc. Both networks are less than 90 deg long; one has a characteristic impedance of 100 ohms, the other of 25 ohms. It is clear that the 100-ohm line has the least reactance variation. Figure 3-35b illustrates the rapid reactance change that may be obtained by selecting an open-circuited line between 90 and 180 deg in length, or a short-circuited line between 180 and 270 deg in length. All three lines are of 100-ohm characteristic impedance, and the lengths are so chosen that the networks all have a reactance of $+j50$ at 500 Mc. In all cases the inductive reactance increases with frequency, but almost any desired slope may be obtained.

The slope of the curve of reactance against frequency for a network that is capacitive behaves in the manner described above, except, of course, that the capacitive reactance of a network decreases with frequency.

3-10. Compensation with the Parallel Network.—In compensating an impedance with a parallel network, several aspects of the impedance curve

must be considered. (1) All the points to be compensated must fall within the boundary circles. (2) The curve must move generally upward (counterclockwise) on an R - X diagram as the frequency increases. The

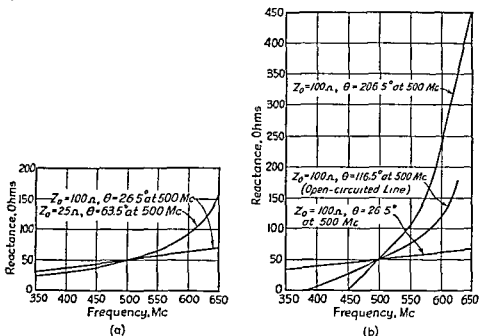


FIG. 3-35—Input reactance vs frequency of several networks

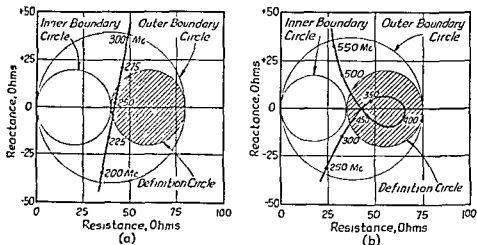


FIG. 3-36—Two types of impedance curves most susceptible to compensation by a parallel network.

two types of impedance curves most readily compensated by a parallel network are shown in Fig. 3-36. The curve in Fig. 3-36b has a loop at approximately midband frequency. The network chosen for this type

of curve will be a short-circuited line $\frac{1}{4}$ wavelength long at midband frequency, or an open-circuited line $\frac{1}{2}$ wavelength long at midband frequency, so that all points at frequencies lower than midband frequency are moved upward (*i.e.*, counterclockwise) and all points at frequencies higher than midband frequency are moved downward (or clockwise). It is obvious that the size of the loop is increased by a parallel network, and thus it must be originally smaller than the definition circle if it is to be compensated in this way.

Example: Let us take the impedance curve in Fig. 3-36b as an example to show all the steps in determining a parallel network. The curve, the definition circle, and the boundary lines are drawn in Fig. 3-36b. The next step is to compute

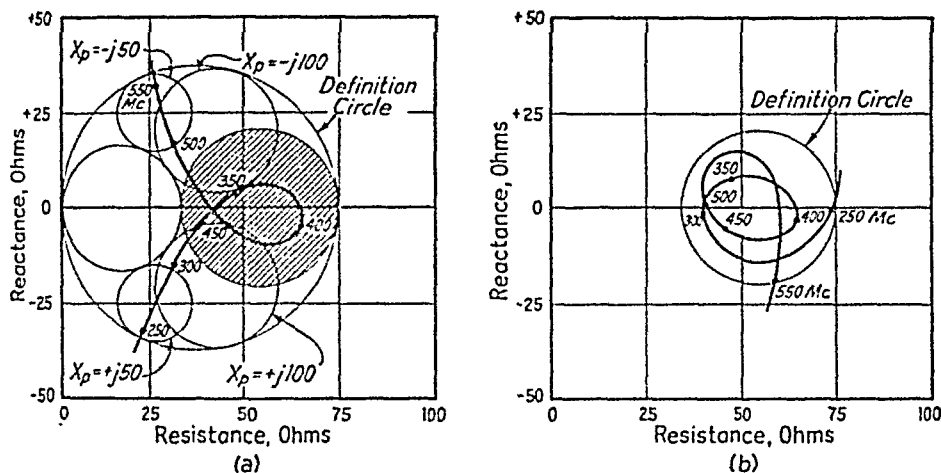


FIG. 3-37.—(a) The impedance curve of Fig. 3-36b plotted with several transformation circles; (b) the input impedance when an open-circuited 100-ohm line 180-deg long at 400-Mc is added in parallel.

several transformation circles, enough so that it is possible to interpolate with reasonable assurance. The definition circle is a 1.5:1-standing-wave-ratio circle for a 50-ohm line, passing through the points $R = 75$, $X = 0$ and $R = 33.3$, $X = 0$. Thus, a is equal to 54.2 and b is equal to 20.8. A quick method of locating a transformation circle is to compute the resistance value of the center

of the circle, which is $R = \frac{a}{1 + (Z_0/X_p)^2}$ where Z_0 is the characteristic impedance of the main transmission line and X_p is the reactance of the parallel network. The resistance centers of the transformation circles for $X_p = \pm j50$ and $X_p = \pm j100$ are $R = 27$ and $R = 43$. It is possible now to locate the center of the corresponding transformation circles with a compass by trial and error, since the center must be on the computed resistance line, equidistant from the two boundary lines. The four transformation circles for $X_p = +j50$, $X_p = +j100$, $X_p = -j50$, and $X_p = -j100$ are drawn in Fig. 3-37a. The impedance curve of Figs. 3-36b and 3-37 is of the type that may be compensated by a short-circuited network either $\frac{1}{4}$ or $\frac{3}{4}$ wavelength long at midband frequency, or an

open-circuited network $\frac{1}{2}$ wavelength long at midband frequency (see Sec. 3-9). The 400-Mc point is defined as the midband point, since it is near the R -axis and since at 400 Mc the curve is moving in a clockwise direction and is therefore critical. The 250-Mc point requires a reactance of $+j50$ ohms for compensation, and the 550-Mc point requires a reactance of $-j50$ ohms. It is now necessary to see whether there is a network that has infinite reactance at 400 Mc, a $+j50$ -ohm reactance at 250 Mc, and a $-j50$ -ohm reactance at 500 Mc.

A short-circuited network 90 deg long at 400 Mc is 56 deg long at 250 Mc and 124 deg long at 550 Mc. Since $X_p = +jZ_0 \tan \theta$, Z_0 must be about 33 ohms if X_p is to be $+j50$ at 250 Mc. A line of 33-ohm characteristic impedance 124 deg long has a reactance of $-j49$ ohms at 550 Mc, which will compensate for the 550-Mc point. The reactance of such a line at 350 Mc is about 165 ohms. From the information on Fig. 3-37a, it seems that the 350-Mc point would probably fall within the $+j165$ circle. The other points on the impedance curve may be considered in a similar manner.

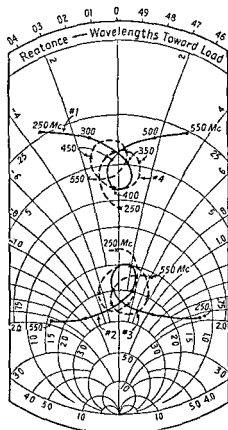
An open-circuited network that is a half-wave long at 400 Mc may also be used. The characteristic impedance of such a network must be higher than that of a quarter-wave short-circuited network and may be determined as shown above.

The input impedance characteristics of a compensated antenna may be computed point by point by using the formula for combining impedances in parallel

$$Z_{in} = \frac{jX_p Z_i}{Z_i + jX_p} \quad (3-25)$$

FIG. 3-38.—The graphical determination on a Smith chart of the input impedance plotted in Fig. 3-37b.

where Z_{in} is the impedance of the compensated antenna, Z_i the impedance of the uncompensated antenna, and jX_p the reactance of the parallel network. Figure 3-37b is a plot of the calculated compensated impedance Z_{in} of the impedance curve in Fig. 3-37a placed in parallel with an open-circuited 100-ohm line 180 deg long at 400 Mc. The input impedance may be determined graphically on a Smith chart, as shown in Fig. 3-38. Curve 1 is the antenna impedance plotted on a Smith chart. Curve 2 is the antenna admittance, each point plotted 180 deg away from the corresponding impedance point. Adding to each point on curve 2 the susceptance value of the parallel network at that frequency gives curve 3, the



input admittance. Curve 4 is the input impedance; hence each point is plotted 180 deg from the corresponding point on curve 3.

3-11. General Considerations in Designing Single-element Broad-band Compensating Networks.—Impedance compensation is used to improve the impedance characteristics of an antenna, *i.e.*, to reduce the magnitude of the reflected wave on the transmission line feeding the antenna. The impedance standard for an antenna designed for a particular purpose may be expressed as the maximum permissible reflection coefficient. A definition circle may be drawn on an R - X diagram enclosing all impedance points that have a reflection coefficient less than the standard. When the measured impedance of an antenna does not fall within the definition circle at all frequencies at which the antenna is to operate, an impedance compensation network may be useful. A great many antennas can be made to operate over a broad band of frequencies by the use of an appropriate compensating

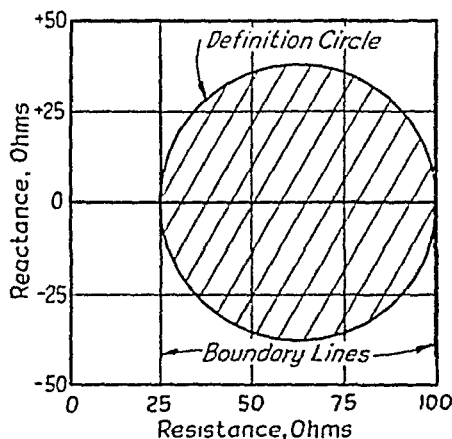


FIG. 3-39.—Series-network boundary lines for a given definition circle.

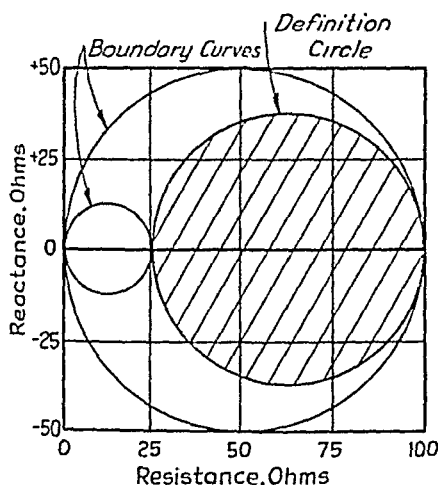


FIG. 3-40.—Parallel-network boundary curves for a given definition circle.

network consisting of only one element, called a *single-element compensating network*.

The procedure in determining the most useful type of compensating network is to plot on an R - X diagram the measured impedance curve and the appropriate definition circle. Usually the most desirable type of network may be selected by inspection when the boundary curves corresponding to each type of network are drawn on the diagram. The boundary curves for a series network and for a parallel network are dependent only on the definition circle. The boundary curves for a series network are two vertical lines passing through the resistance extremes of the definition circle, the $R = Z_0 \times \sigma$ line and the $R = Z_0/\sigma$ line, where σ is the voltage standing-wave ratio (Fig. 3-39). The boundary curves

for a parallel network are two circles, passing through the origin, centered on the R -axis and passing, respectively, through the two resistance extremes of the definition circle, as shown in Fig. 3-40. The boundary circles for a line transformer are dependent not only on the definition circle, but also on the characteristic impedance Z_0' of the line trans-

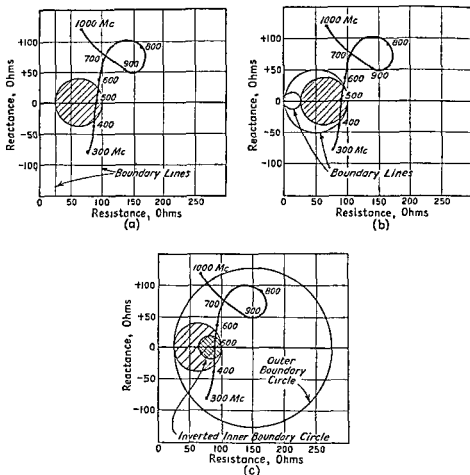


FIG. 3-41—An impedance curve and definition circle plotted with (a) series-network boundary lines, (b) parallel-network boundary circles, and (c) line-transformer boundary circles

former. The optimum line-transformer boundary circles for a specified definition circle and particular antenna impedance curve may be constructed by inspection, as discussed in Sec. 3-5 and illustrated in Fig. 3-16.

Example: Figure 3-41 is the impedance curve of Fig. 3-4, with the series network boundary lines in *a*, the parallel network boundary lines in *b*, and the line-transformer boundary lines in *c*. The definition circle to which this impedance

curve is to be transformed is a circle of 0.33 reflection coefficient on a 50-ohm line, passing through the points $R = 25$, $X = 0$ and $R = 100$, $X = 0$. Since in both cases the major part of the curve lies outside the boundary curves, it is obvious that neither a series network nor a parallel network would be effective in improving the impedance characteristics of this curve. A line transformer, however, might be effective. The boundary curves for the line transformer are constructed in the following way. The outer boundary circle is drawn tangent to the definition circle at $R = 25$ with a radius arbitrarily chosen large enough so that the 300-Mc point is well within the circle. The high resistance R -intercept of this circle is at $R = 276$. The inner boundary circle intercepts the R -axis at two points, at the resistance maximum of the definition circle and at the point where R is equal to the outer boundary-circle high-resistance intercept times the ratio of the definition-circle intercepts; in this case at $276 \times 25/100$ or at $R = 69$. Thus the inner boundary circle lies within the definition circle and is inverted (see Sec. 3-5). The characteristic impedance of a line transformer having the boundary circles shown in Fig. 3-41c is 83 ohms, the geometric mean between 25 and 276, the R -intercepts of the outer boundary circle, or 69 and 100, the R -intercepts of the inner boundary circle.

In order to determine the optimum length of the line transformer, it is necessary to construct transformation circles for several values of θ . The $\theta = 90$ -deg circle is tangent to the outer boundary circle at its high-resistance R -intercept and tangent to the inverted inner boundary circle at its low-resistance R -intercept, as shown in Fig. 3-42. The $\theta = 45$ -deg and $\theta = 135$ -deg circles are tangent to the inner and outer boundary circles at the points where they are crossed by a circle of radius Z_0' , in this case 83, centered at the origin. Since the inner boundary circle is inverted, the transformation circles all include the inner boundary circle. The $\theta = 45$ -deg and $\theta = 135$ -deg transformation circles are drawn on Fig. 3-42. The 800-Mc point is the most critical, since at that frequency the

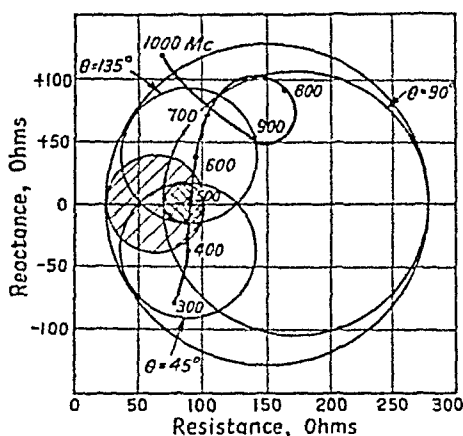


FIG. 3-42.—Boundary curves and transformation circles required to estimate the length of the line transformer needed to compensate the impedance curve.

TABLE 3-3	
Frequency, Mc	θ , Deg
300	39
400	53
500	66
600	79
700	92
800	105
900	118
1000	132

impedance curve is moving clockwise (see Sec. 3-3). Estimating the position of the various θ circles near the 800-Mc point, one could assume that the 800-Mc point is closest to the center of the $\theta = 105$ -deg circle. At the other frequencies on the impedance curve, θ has the values shown in Table 3-3. By estimating the position of each θ circle, and seeing whether or not the corresponding imped-

ance point falls within it, the effectiveness of the line transformer may be determined. The 300-Mc point is probably about on the edge of the $\theta = 39$ -deg circle; the 400-Mc point is certainly within the $\theta = 53$ -deg circle; the 500-Mc point is within the inverted inner boundary circle and, consequently, cannot be transformed out of the definition circle; the 600-Mc point is very probably within the $\theta = 79$ -deg circle, the 700-Mc point must lie within the $\theta = 92$ -deg circle, since it is in both the $\theta = 90$ -deg circle and the $\theta = 135$ -deg circle; the 800-Mc point is, of course, within the $\theta = 105$ -deg circle, and the 900-Mc point must lie within the $\theta = 118$ -deg circle, since it lies within both the $\theta = 90$ -deg and $\theta = 135$ -deg circles. Therefore, a line transformer of 83-ohm characteristic impedance, 105 deg long at 800 Mc will transform the impedance curve to the definition circle, the only questionable points being the 300-Mc point and the 600-Mc point. Each point may be transformed on a Smith chart to determine the input impedance of the compensated antenna. The computed input impedance is plotted on Fig. 3-4b.

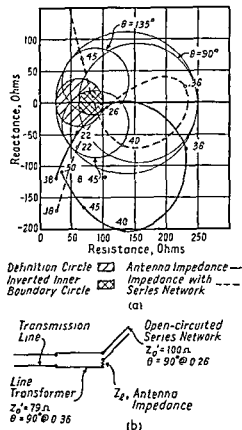


FIG. 3-43.—An example of two-element impedance compensation: (a) the transformation of the solid-line antenna-impedance curve by an open-circuited series network of 100-ohm characteristic impedance 90 deg long at $f = 0.26$ to the dotted-line curve which can be transformed by a line transformer of 79-ohm characteristic impedance 90 deg long at $f = 0.36$, (b) a diagrammatic representation of the compensating system.

to compensate an impedance curve that cannot be compensated by a single element. The theory of a multielement compensating network is to use the first element to transform the impedance curve to a position from which it can either be compensated by the second element, or be transformed by the second element to a position from which it can be compensated by a third element, etc.

3-12. Multielement Compensating Networks.—It is possible to use several networks together

Example: An impedance curve compensated by a two-element network is shown in Fig. 3-43a. The principles used in selecting the proper network types are similar to those used for single-element networks.

The solid-line impedance curve of Fig. 3-43a (plotted with relative frequency) cannot be compensated to the definition circle by any single-element network, since it lies outside the boundary lines for any type of network. However, it seems possible by inspection that the curve could be transformed so that it would lie within the boundary curves of a line transformer by a series network that is inductive at the higher frequencies. The boundary curves and the transformation circles for a line transformer of 79-ohm characteristic impedance are drawn on Fig. 3-43a. The outer boundary line was chosen to include a resistance equal to the highest resistance at any point on the impedance curve. Under the assumption that this point, which in Fig. 3-43a is the 0.36-impedance point, is to be transformed by the series network to a position near the R -axis, a line transformer 90 deg long may be used to transform this point to the definition circle.

A line transformer 90 deg long at a relative frequency 0.36 has the lengths given in Table 3-4 at the other frequencies. Therefore, in order that the line-transformer transformation be successful, the series network must transform each

TABLE 3-4

Relative frequency	θ	X_s limits	X_s of suggested network
0.18	45	Not possible	
0.22	55	$-j30$ to $+j40$	$-j25$
0.26	65	$-j80$ to $+j25$	0
0.36	90	$+j25$ to $+j110$	$+j70$
0.40	100	$+j130$ to $+j280$	$+j146$
0.45	110	$+j155$ to $+j250$	$+j244$
0.50	125	Not possible	

impedance point to a position somewhere within the corresponding θ circle. Table 3-4 includes the approximate values of X_s required to transform each point to the corresponding θ circle. It is now necessary only to determine a series network that has a reactance value at each relative frequency within the limits given in Table 3-4. An open-circuited network 90 deg long at the relative frequency 0.26 with a characteristic impedance of 100 ohms has the reactance values shown in Table 3-4. The dashed-line impedance curve on Fig. 3-43a is the impedance resulting from the addition of this series network to the antenna impedance. It is obvious that a line transformer of 79-ohm characteristic impedance 90 deg long at the relative frequency 0.36 will be successful in transforming the dashed-line impedance curve. The resulting system is shown in Fig. 3-43b.

Other combinations of networks may be used for impedance compensation, following the principles discussed above.

3-13. Baluns: General.—The problem of transmitting energy from coaxial or unbalanced transmission-line systems to dual or balanced trans-

mission-line systems, or vice versa, occurs frequently in broad-band antenna work. Balanced antennas in general are better suited to broad-band applications than are unbalanced antennas. Since antennas are almost always fed by coaxial lines, there is need for a device to convert the unbalanced voltage of the coaxial cable to the balanced voltage required by the antenna. This balance-to-unbalance transformer is commonly called a *balun*. In the following sections some of the forms that the balun may take will be discussed.

3-14. Type-I Balun.—A type-I balun, sometimes called a *bazooka*, is

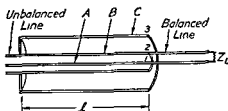


FIG. 3-44 — Type-I balun, sometimes called a "bazooka."

shown in Fig. 3-44. In this balun the balanced pair of conductors is tied directly to the inner and outer conductors of the coaxial line at points 1 and 2. In addition, a skirt *C*, placed around the coaxial line, extends $\frac{1}{4}$ wavelength from its junction with the balanced line.

Thus, the outer conductor *B* of the coaxial line and the skirt *C* form a second coaxial line the impedance of which, as measured between 2 and 3, is

$$Z_{23} = Z_{0BC} \tan \left(\frac{2\pi l}{\lambda} \right) \quad (3-26)$$

where Z_{23} is the impedance seen between 2 and 3, Z_{0BC} the characteristic impedance of the coaxial line consisting of outer conductor *B* and skirt *C*, l the length of the skirt *C*, and λ the wavelength at the operating frequency.

When $l = \lambda/4$, the impedance Z_{23} is very high (limited by the Q of the section) and there is no shunt path to ground from 2 to 3 and consequently no division of current at junction 2. Thus, both conductors of the dual line carry equal currents.

There are, however, paths to ground from each of the dual conductors because of the distributed capacitance and conductance. Since these paths are presumably equal and since the currents in each of the dual conductors are equal and 180 deg out of phase, the voltages to ground will be equal and 180 deg out of phase and the dual line will be balanced.

If the length of the skirt is not $\frac{1}{4}$ wavelength, however, Z_{23} will not be a high impedance, and it is apparent from the equivalent circuit shown in Fig. 3-45 that an unbalanced path will exist between 2 and ground. Therefore, the impedance to ground from 1 will be significantly

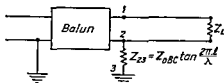


FIG. 3-45.—Equivalent circuit for type-I balun.

different from the impedance to ground from 2 and there will be a division of current at junction 2. The voltages as measured from 1 to ground and from 2 to ground will no longer be equal or 180 deg out of phase, and the dual line will be unbalanced. The band width of this type of balun is therefore very limited, and the balun is suitable for use only in installations where the equipment is to be operated over a band not more than about 10 per cent wide.

3-15. Type-II Balun.—The band width of the bazooka may be increased by utilizing another section of transmission line having exactly the same characteristics as the section formed by outer conductor *B* and skirt *C* of the bazooka. As shown in Fig. 3-46, this added section consists of *B'* and *C'*, with *B'* connected to junction 1. Thus, whatever the impedance between 2 and ground due to *B* and *C*, it will have at all frequencies an exact counterpart between 1 and ground due to *B'* and *C'*. This is illustrated in the equivalent circuit (Fig. 3-47). Because the currents entering junctions 1 and 2 from the coaxial line are equal and 180 deg out of phase, and because the impedance between junctions 1 and 2 to ground are equal at all frequencies, the dual line will be balanced at all frequencies. At the frequency at which the lengths of *B* and *B'* are exactly $\frac{1}{4}$ wavelength, the impedance presented to the coaxial line will be only the input impedance of the dual line since the impedance of the sections of the balun is infinite. At any other frequency the balun will shunt the dual line with the impedance equal to $j2Z_{0BC} \tan \left(2 \frac{\pi l}{\lambda} \right)$,

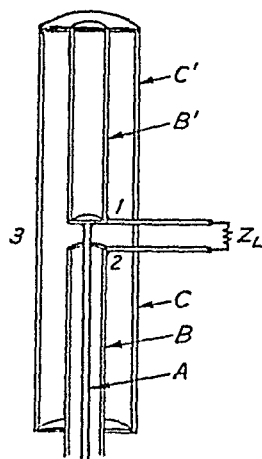


FIG. 3-46.—Type-II balun.

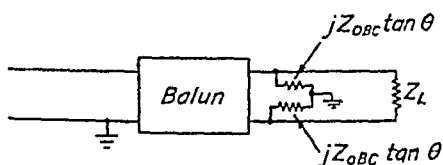


Fig. 3-47.—Equivalent circuit for type-II balun.

where Z_{0BC} is the characteristic impedance of each leg of the balun.

If the gap between *B* and *B'* is made large compared with the shortest wavelength at which the balun is to be used, the section of inner conductor of the coaxial cable will place inductance in series with

one side and some unbalance will be present. This gap should therefore be made as small as is consistent with consideration of voltage arc-over and lumped capacitance between 1 and 2. Practical values that have been used range from $\frac{1}{8}$ to $\frac{1}{4}$ in.

Generally, the characteristic impedance of the dual line is greater than the impedance of the coaxial line. Advantage can be taken of the

shunt reactance presented by the balun by utilizing it as part of the impedance-matching network. An example in Sec. 3-17 illustrates this.

3-16. Type-III Balun.—The type III operates in the same manner as the type-II balun except that section B' is placed beside section B so as to form a dual line within the skirt C , as shown in Fig. 3-48. As pointed out in the preceding section, the length d of the inner conductor of the coaxial line must be kept short to minimize unbalance. Type-III construction does not lend itself to keeping this length short since, as can be seen from Fig. 3-48, the inner conductor has to traverse the distance between the sections B and B' and this distance is determined by the characteristic impedance desired in the balun.

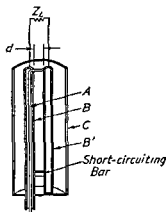


FIG. 3-48—Type-III balun.

The principal advantage of a type-III balun is that its length need be only half that of a type-II balun. Furthermore, a short-circuiting bar can be placed as shown to permit adjustment of the length of the balun for operation at different frequencies.

3-17. Use of the Balun as Part of an Impedance-matching Network.—In the preceding section it was pointed out that, while the dual line is balanced at all frequencies, the balun sections will place a reactance in

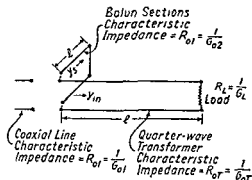


FIG. 3-49.—Schematic diagram of balun used with quarter-wave transformer for broad-band impedance matching.

shunt with the dual line at the junction with the balun. This shunt reactance can be incorporated as part of the impedance-matching network between load and coaxial line, as illustrated in the following example

Example: Let it be assumed that a coaxial cable having a characteristic impedance R_{01} is to be used with an ideal balanced load of resistance $R_L = 2R_{01}$. An ideal load is taken to mean one that does not change its impedance with frequency.

The maximum degree of mismatch that can be tolerated is assumed to be that which gives a 1.5:1 voltage standing-wave ratio. The impedance-matching network is to consist of the balun sections and a quarter-wave transformer, as shown in Fig. 3-49. It is required to find the characteristic impedance of the quarter-wave transformer and of the balun sections to give maximum band width, and to find the frequency range over which the required impedance match can be obtained.

Since the impedances are in parallel, it will be more convenient to work in terms of admittance rather than impedance.

Then

$$G_L = \frac{1}{R_L} \quad (3-27)$$

$$G_{01} = \frac{1}{R_{01}} \quad (3-28)$$

$$2G_L = G_{01} \quad (3-29)$$

First, let us plot, in terms of G_{01} , the characteristic conductance G_{01} of the coaxial cable and the conductance G_L of the load on the admittance chart of Fig. 3-50. The locus of all admittances that will produce a 1.5:1 voltage-standing-wave ratio on the coaxial cable when used to terminate the cable is the circle given in Fig. 3-50 as curve 1. Admittances lying within this circle will produce standing-wave ratios of less than 1.5:1. Therefore, it is necessary to use a device that, when terminated in the load admittance G_L , will have an input admittance lying within this circle over as much of a frequency band as possible. With this object in mind let us choose the quarter-wave transformer so that it is $\frac{1}{4}$ wavelength long at the center frequency, and so that it will transform G_L to the maximum conductance point on the 1.5:1 circle. Thus the characteristic conductance $G_{0T} = 1/R_{0T}$ of the transformer becomes

$$G_{0T}^2 = (1.5G_{01})G_L \quad (3-30)$$

$$G_{0T} = \sqrt{1.5G_{01}G_L} \quad (3-31)$$

Substituting the value of G_L from Eq. (3-29), we get

$$G_{0T} = \sqrt{\frac{1.5}{2}} G_{01}^2 = 0.866G_{01} \quad (3-32)$$

If we consider only the portion of the system consisting of the load G_L and the quarter-wave transformer, the input admittance to the transformer will vary with frequency between the values G_L and $1.5G_{01}$ along the circle shown as curve 2

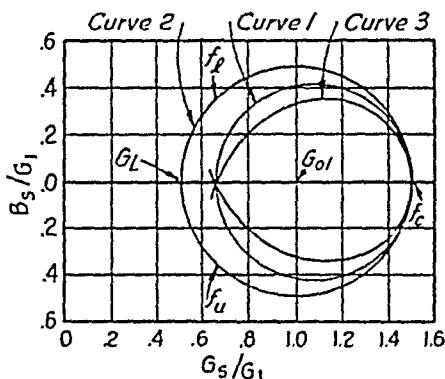


FIG. 3-50.—Admittance chart for the circuit of Fig. 3-49.

on Fig. 3-50. The equation of this circle may be found from the transmission-line equations. Calling the input admittance to the transformer Y_{in} , we may write

$$\frac{Y_{in}}{G_{or}} = \frac{\frac{G_{or}}{G_L} - j \tan \theta + j \left(\frac{G_{or}}{G_L}\right)^2 \tan \theta + \frac{G_{or}}{G_L} \tan^2 \theta}{\left(\frac{G_{or}}{G_L}\right)^2 + \tan^2 \theta} \quad (3-33)$$

$$\frac{G_{in}}{G_{or}} = \frac{\frac{G_{or}}{G_L} (1 + \tan^2 \theta)}{\left(\frac{G_{or}}{G_L}\right)^2 + \tan^2 \theta} \quad (3-34)$$

$$j \frac{B_{in}}{G_{or}} = \frac{j \tan \theta \left[\left(\frac{G_{or}}{G_L}\right)^2 - 1 \right]}{\left(\frac{G_{or}}{G_L}\right)^2 + \tan^2 \theta} \quad (3-35)$$

Substituting from Eqs. (3-26) and (3-29), we get

$$\frac{G_{in}}{G_{o1}} = \frac{1.5(1 + \tan^2 \theta)}{3 + \tan^2 \theta} \quad (3-36)$$

$$j \frac{B_{in}}{G_{o1}} = j \frac{\sqrt{3} \tan \theta}{3 + \tan^2 \theta} \quad (3-37)$$

Equations (3-36) and (3-37) are the equations of curve 2.

The quarter-wave transformer has effected the required degree of match at f_c only, where the two circles are tangent. At every other frequency the admittance presented to the coaxial cable lies outside the 1.5:1 circle. However, for those frequencies where the real part of Y_{in} lies between the values $0.66G_{o1}$ and $1.5G_{o1}$ it is possible to collapse curve 2 to within the 1.5:1 circle by tuning out part of the reactive component of Y_{in} . The reactance of the balun sections can be used to accomplish this tuning if the characteristic impedance of the balun is properly chosen.

The frequency limits at which G_{in} is equal to $0.66G_{o1}$ are found by substituting this value in Eq. (3-36). Then

$$\frac{G_{in}}{G_{o1}} = 0.66 = \frac{1.5(1 + \tan^2 \theta)}{3 + \tan^2 \theta} \quad (3-38)$$

$$\tan \theta = \pm \sqrt{0.6} = \pm 0.775 \quad (3-39)$$

$$\theta = 37.8^\circ, 142.2^\circ \quad (3-40)$$

At the center frequency f_c the electrical length of the transformer was made equal to $\frac{1}{4}$ wavelength, or 90 deg. Since electrical length is directly proportional to frequency, we may easily find the frequency limits corresponding to the value of θ in Eq. (3-40). Thus the lower limit is $f_l = 37.8f_c/90 = 0.42f_c$, and the upper limit is $f_u = 142.2f_c/90 = 1.58f_c$.

The susceptive component of admittance B_{in} presented by the transformer at these frequency limits is found from Eq. (3-37) to be $\pm j0.373G_{o1}$.

Up to this point the characteristics of the balun have not been mentioned. If the balun sections are made to be $\frac{1}{4}$ wavelength long at the center frequency f_c , then at this frequency the balun will shunt the quarter-wave transformer with a very low admittance. At frequencies other than f_c , the balun will act as a shunt to the dual line at the junction and will have an admittance Y_s equal to

$$Y_s = -jG_{02} \cot \theta \quad (3-41)$$

where G_{02} is the characteristic conductance of the balun. In the case of the type-II balun, G_{02} is half the characteristic conductance of each section of the balun. If G_{02} is chosen so that at f_l and f_u the value of Y_s will be exactly equal to, but of different sign from, B_{in} , then at these frequencies the total admittance seen by the coaxial cable will be $0.66G_{01}$ and the effect of the balun stubs will be to transform curve 2 to curve 3.

The value of G_{02} that will accomplish this is found by equating B_{in} in Eq. (3-37) to $-Y_s$ in Eq. (3-41) for $\tan \theta = \pm 0.775$. Then

$$\frac{G_{01} \sqrt{3} \tan \theta}{3 + \tan^2 \theta} = G_{02} \cot \theta \quad (3-42)$$

and

$$G_{02} = 0.288G_{01} \quad (3-43)$$

If at each frequency the net admittance presented to the coaxial line by balun and transformer is plotted, curve 3 will be the result. Thus it can be seen that for all frequencies between f_l and f_u the admittance terminating the coaxial cable lies within the 1.5:1 circle and the balun serves usefully as a broad-band device over a 3.76-to-1 frequency range. This arrangement can work only over a limited frequency range because at $2f_c$ the balun will short-circuit the dual line.

3-18. Methods of Checking Balun Performance.—The chief item of concern in the performance of a balun is the degree of balance that exists on the output dual line. This may be determined by measuring with a suitable voltmeter the equality of the voltages from each of the dual conductors to ground. The voltmeter must have a very high input impedance so that it will not introduce any unbalance of its own. Figure 3-51 shows a type of voltmeter that has been used for this purpose. Basically, it is a quarter-wave coaxial line short-circuited at one end by a high capacitance. A crystal and a d-c meter serve to detect the short-circuit current, which is directly proportional to the voltage at the input. If the length of the coaxial-line voltmeter is exactly $\frac{1}{4}$ wavelength, the input impedance to the voltmeter will be very high.

Since the voltmeter gives a reading proportional to amplitude, this

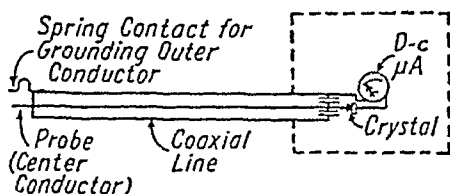


FIG. 3-51.—Schematic diagram of r-f voltmeter.

method will not detect *phase unbalance*, i.e., unbalance in which the peak amplitudes of the voltages on the two lines are equal but do not occur 180 deg apart in time. A better method of detecting unbalance, and one that does not have this shortcoming, is illustrated in Fig. 3-52. Here a quarter-wave stub is shunted across the dual line. Since the stub is

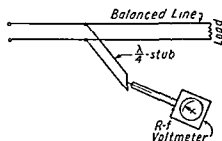


FIG. 3-52 — Schematic diagram of a setup for measuring unbalanced voltage on a balanced line

short-circuited, it will present infinite impedance to the balanced voltage. For the unbalanced voltages, however, the quarter-wave stub is open-circuited at its end so that the points on the dual line at which the stub is attached are short-circuited for unbalanced currents. Hence, there will be some voltage from the end of the stub to ground if any unbalance exists on the line.

CHAPTER 4

CONE AND CYLINDER ANTENNAS

By A. DORNE

The cone antenna, when its side is several wavelengths long and its included angle is fairly large, is usually regarded as a type of horn. In fact, when two such cones are placed with their apexes together, the term *biconical horn* is accepted as standard. On the other hand, a cone antenna that has an included angle of only a few degrees and a side $\frac{1}{4}$ wavelength long is much more naturally regarded as a stub. From a theoretical point of view any distinction between the two appears somewhat arbitrary and unnecessary. But from the point of view of uses and applications a distinction between these two types may be quite meaningful. It is on this basis that the material for this section has been selected. Thus, under the classification *cone and cylinder antennas* is included information about antennas of these types that is of interest when they are used as the broad-band analogues of the long thin dipoles or vertical masts of ordinary narrow-band practice.

4-1. General Considerations:

Impedance.—As discussed in Chap. 1, an antenna of any type may be regarded as a transformer between guided and free waves. Since, in the case of conical antennas, this is a particularly useful point of view, an example is given to illustrate the sense in which the word is used.

Consider a junction between two transmission lines, one consisting of a coaxial line and the other of a section of guide of circular cross section, as shown in Fig. 4-1. The section between planes *AA* and *BB* consists of a section of transmission line that functions as a transformer between the coaxial line and the guide. If the diameter of the guide is now increased until it becomes infinite, the transformer section remains a transformer, but the guide has become free space, so that the transformation is between the guided waves in the coaxial line and the free waves in

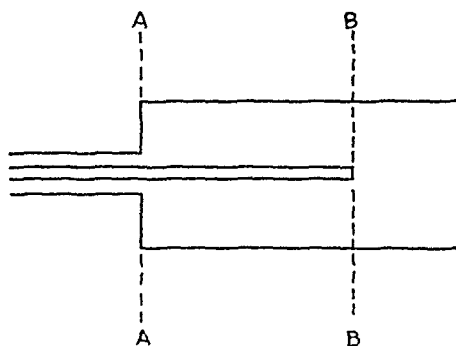


FIG. 4-1.—Transformer from coaxial transmission line to cylindrical hollow waveguide.

space. As with any other transformer, if it is reflectionless or nearly so, it will be broad band, while if it produces large reflections, it may still be a satisfactory transformer but only over a narrow frequency band (or bands) within which the various reflections cancel each other. In general, as is pointed out in other chapters, a gradual transformation between two types of waves is likely to be reflectionless, whereas an abrupt change is not.

If we consider the behavior of a conical antenna, we shall see that it is reasonable that it should produce only small reflections once it becomes reasonably large in terms of wavelengths. Schelkunoff¹ has devised

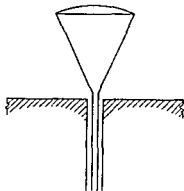


FIG. 4-2.—Conical antenna fed by coaxial transmission line

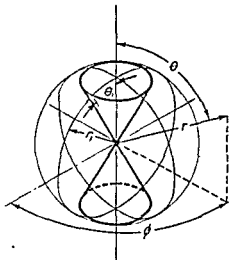


FIG. 4-3.—Biconical antenna.

a method for computing the behavior of the conical antenna in terms of electromagnetic theory. The mathematics becomes quite involved, however, and although this discussion follows his basic analysis of the problem, no quantitative solution is attempted.

In terms of a conical antenna the situation of Fig. 4-1 becomes that of Fig. 4-2, in which are shown the coaxial feed line, a conical antenna, and an infinite ground plane. The problem is slightly simplified if we remove the ground plane and replace the image of the antenna by a real antenna, which is symmetrical with the first. We then have a balanced biconical radiator energized by a voltage applied between the apexes of the two cones, as shown in Fig. 4-3. In this figure the cones are equal and collinear and are bounded at the ends by the spherical surface that has its center at the center of the antenna. The region within the sphere

¹ SCHELKUNOFF, S. A., "Electromagnetic Waves," D. Van Nostrand Company, Inc., New York, 1943.

may be regarded as the section of transmission line that comprises the transformer.

This is a natural boundary to choose because the waves that can exist in this system possess spherical fronts, and if a particular wave extended only as far as the end of the antenna, it would, because it was spherical, extend as far as the sphere bounding the antennas. Accordingly, the spherical surface constitutes a boundary between free space and the transformer. There are three types of waves that can exist in the system. One of these is a *TEM* wave. Its existence depends on the presence of conductors to bound its field, and it can therefore exist only within the boundary sphere. It is similar to the type that is normally present in a coaxial line in that neither its electric nor magnetic field possesses a component in the direction of propagation of the wave. Furthermore, lines of electric force extend from one cone to the other, in correspondence with the lines of electric force in a coaxial transmission line that extend from one conductor to the other. The expressions for the field are given in Eqs. (4-3) and (4-4), and a cross-sectional view of the lines of electric force is shown in Fig. 4-4.

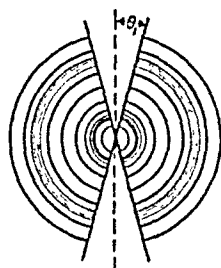
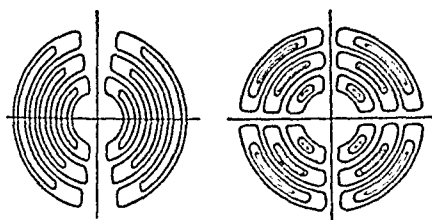


FIG. 4-4.—Lines of electric force for *TEM* wave in the presence of two collinear conical conductors. (From S. A. Schelkunoff, "Electromagnetic Waves," D. Van Nostrand Company, Inc., New York, 1943, by permission.)

Outside the sphere an infinite family of waves of the *TM* type can exist. The summation of these constitutes the free wave traveling through space away from the antenna. Like other *TM* waves, their magnetic fields have no components in the direction of propagation, although their electric

fields do. Figure 4-5 shows a cross-sectional view of the lines of electric force for the first- and second-order *TM* waves in free space. The expressions for the field components are given in Eqs. (4-5) to (4-10). The third possible type of wave also consists of an infinite family of *TM* waves. These, however, exist within the sphere and are modified by the presence of the cones. Figure 4-6 shows cross-sectional views



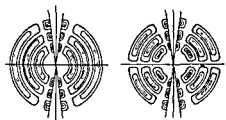
First-order *TM* Wave Second-order *TM* Wave

FIG. 4-5.—Lines of electric force for first and second order *TM* waves in free space. (From S. A. Schelkunoff, "Electromagnetic Waves," D. Van Nostrand Company, Inc., New York, 1943, by permission.)

of the lines of electric force for the first- and second-order waves of this type.

The method of solution of the problem consists of obtaining relative values for these various waves such that the resultant field at the surface of the sphere due to the waves inside it matches the field due to the waves

outside it. Schelkunoff points out an interesting aspect of the problem, which greatly facilitates relating the results thus obtained to the input impedance of the antenna. This is, that the *TM* waves vanish as the apexes of the cones are approached and, therefore, the current and the voltage at the input of the antenna are determined completely by the *TEM* wave. It can readily be shown that the correspondence between this *TEM* wave and the fundamental wave of an ordinary coaxial line is closer than was indicated above. Specifically, the following properties are common to both (on the assumption of no loss in the dielectric or the conductors):



First-order *TM* Wave Second order *TM* Wave

FIG. 4-6.—Lines of electric force for first and second order *TM* waves in the presence of two collinear conical conductors (From S. A. Schelkunoff, "Electromagnetic Waves," D. Van Nostrand Company, Inc., New York, 1943, by permission)

The velocity of propagation is independent of frequency and is equal to the velocity in the medium (in this case free space). If there is just one wave traveling along either the coaxial line or the transmission line composed of the two cones (as in the case of a line of infinite length), the current and voltage between conductors is the same at all points along the line. And finally, the characteristic impedance Z_0 is the same at all points along the line, is independent of frequency, and is given for the biconical transmission line by the expression

$$Z_0 = \frac{1}{2} \sqrt{\frac{\mu}{\epsilon}} \frac{1}{\pi} \ln \cot \frac{\theta_1}{2} \quad (4-1)$$

in which θ_1 is the angle between the side of one of the cones and its axis, μ the permeability (magnetic inductive capacity) of the medium, and ϵ its dielectric constant (electric inductive capacity).

It therefore appears that the input impedance of the antenna can be determined from the characteristic impedance of the line composed of the two cones and the relative amplitude and phase of the transmitted and reflected *TEM* waves. This is the first reference to a reflected *TEM* wave; it is apparent, however, that it not only can, but must, exist, for otherwise the input impedance of the antenna would simply be the Z_0 of the line represented by the antenna regardless of frequency. This could be true only if the cones were infinitely long. It will also be realized that the transmitted and reflected *TEM* waves cannot be of equal intensity, for if they were, no energy would be transmitted by the antenna. Accordingly, if the reflection coefficient of the *TEM* wave at the boundary sphere were known, the input impedance of the antenna could be computed by means of the ordinary transmission-line equation

$$Z_{\text{input}} = Z_0 \frac{1 - \rho e^{-2i\beta r_1}}{1 + \rho e^{-2i\beta r_1}} \quad (4-2)$$

in which Z_0 is the characteristic impedance of the transmission line composed of the cones, ρ the complex current reflection coefficient, r_1 the length of the cone side, and β is $2\pi/\lambda$.

Finding the value of this reflection coefficient is not simple. It is readily possible, however, to deduce some qualitative information about its behavior by considering the relation at the boundary sphere between the *TEM* wave and the other waves in the system. The following discussion is concerned with an approach of this kind.

It is reasonable to suppose that, if at the spherical surface a summation of free-space *TM* waves can approximately duplicate the field due to the incident *TEM* wave, then there is no reason for a large reflected *TEM* wave to exist. Examination of the expressions for the components of the various waves shows that we may expect this condition to obtain when the radius of the sphere is large in terms of wavelength. For the *TEM* wave the instantaneous expressions for the field components¹ are

$$\mathcal{H}_\phi = \frac{e^{-i\beta r}}{r \sin \theta} \quad (4-3)$$

$$\mathcal{E}_\theta = \sqrt{\frac{\mu}{\epsilon}} \frac{e^{-i\beta r}}{r \sin \theta} \quad (4-4)$$

In these expressions, μ , ϵ , and β have the meanings assigned in connection with Eqs. (4-1) and (4-2). The definitions of r and θ are given in Fig. 4-3, from which it can be seen that r is the distance from the origin and θ the angle from the vertical axis. Finally, \mathcal{H}_ϕ is the magnetic-field component in the direction of increasing ϕ and \mathcal{E}_θ the electric-field component in the direction of increasing θ . Similarly, \mathcal{E}_r [which appears in Eqs. (4-7) and (4-10)] is the electric-field component in the direction of increasing r .

For the lowest order *TM* wave in free space the only components² are

$$\mathcal{H}_\phi = \frac{e^{-i\beta r}}{r} \left(1 + \frac{1}{j\beta r} \right) \sin \theta \quad (4-5)$$

$$\mathcal{E}_\theta = \sqrt{\frac{\mu}{\epsilon}} \frac{e^{-i\beta r}}{r} \left[1 + \frac{1}{j\beta r} + \frac{1}{(j\beta r)^2} \right] \sin \theta \quad (4-6)$$

$$\mathcal{E}_r = \frac{2e^{-i\beta r}}{j\omega\epsilon r^2} \left(1 + \frac{1}{j\beta r} \right) \cos \theta \quad (4-7)$$

¹ See, for instance, SLATER, J. C., "Microwave Transmission," Chap. V, Sec. 28, McGraw-Hill Book Company, Inc., New York, 1942.

² The expressions for the field components of the *TM* waves given in Eqs. (4-5) to (4-10) may be obtained by rewriting the expressions given by Schelkunoff, *op. cit.*, Chap. XI, Sec. 1.

and for the second-order *TM* wave in free space the only components are

$$\mathcal{H}_\theta = \frac{3}{2} \frac{e^{-j\beta r}}{2r} \left[1 + \frac{3}{j\beta r} + \frac{3}{(j\beta r)^2} \right] \sin 2\theta \quad (4-8)$$

$$\mathcal{E}_\theta = \frac{3}{2} \sqrt{\frac{\mu}{\epsilon}} \frac{e^{-j\beta r}}{r} \left[1 + \frac{3}{j\beta r} + \frac{6}{(j\beta r)^2} + \frac{6}{(j\beta r)^3} \right] \sin 2\theta \quad (4-9)$$

$$\mathcal{E}_r = \frac{3e^{-j\beta r}}{j\omega\epsilon r^2} \left[1 + \frac{3}{j\beta r} + \frac{3}{(j\beta r)^2} \right] (3 \cos^2 \theta - 1) \quad (4-10)$$

If r is sufficiently large, Eq. (4-6) becomes

$$\mathcal{E}_\theta = \sqrt{\frac{\mu}{\epsilon}} \frac{e^{-j\beta r}}{r} \sin \theta$$

and Eq. (4-9) becomes

$$\mathcal{E}_\theta = \frac{3}{2} \sqrt{\frac{\mu}{\epsilon}} \frac{e^{-j\beta r}}{r} \sin 2\theta$$

The expressions for higher order waves are similar and have trigonometric terms that involve higher and higher harmonics of θ . For example, for the wave of the third order, the trigonometric term is $5 \sin 3\theta \sin \theta$. Consequently, the terms are very similar to the terms of a Fourier series, and a suitable summation of them will approximate the distribution of the *TEM* wave that varies as $1/\sin \theta$. It can readily be shown from the above expressions that (when r is sufficiently large) if the boundary conditions are satisfied for \mathcal{E}_θ they are also satisfied for \mathcal{H}_θ because the radial impedance, i.e., the ratio of $\mathcal{E}_\theta/\mathcal{H}_\theta$ for the various waves, becomes simply $\sqrt{\mu/\epsilon}$ in all cases. Further, \mathcal{E}_r diminishes rapidly with increase of r since it varies as $1/r^2$. Thus we are led to the conclusion that as the cone becomes large in terms of wavelengths the reflected *TEM* wave becomes small and, therefore, the input impedance of the antenna approaches the characteristic impedance of the cone transmission line.

Now let us consider what happens when r is considerably smaller than a quarter wave. It can be seen from Eqs. (4-5), (4-6), (4-8), and (4-9) that for each of the two lowest order free-space *TM* waves, at the boundary sphere, the ratio of \mathcal{E}_θ to \mathcal{H}_θ at any point is much greater than the ratio of \mathcal{E}_θ to \mathcal{H}_θ of a single *TEM* wave and consists primarily of a negative reactance term. It can be shown that this is also true for all higher order free-space *TM* waves. Therefore, if the *TEM* wave is to satisfy the boundary conditions to even a first approximation, it too must have a ratio of \mathcal{E}_θ to \mathcal{H}_θ that consists primarily of a negative reactance term that is much larger than $\mathcal{E}_\theta/\mathcal{H}_\theta$ for either the incident or reflected wave. That is,

$$\frac{\mathcal{E}_\theta \text{ incident} + \mathcal{E}_\theta \text{ reflected}}{\mathcal{H}_\theta \text{ incident} - \mathcal{H}_\theta \text{ reflected}} = c - j b \quad (4-11)$$

where $b \gg a$ and b is also $\gg \epsilon_0/\mathcal{C}_\phi$ for either traveling wave. Since this is a *TEM* mode, it can be shown that if Eq. (4-11) applies,

$$\frac{V \text{ incident} + V \text{ reflected}}{i \text{ incident} - i \text{ reflected}} = R - jX \quad (4-12)$$

where

$$X \gg R \quad \text{and} \quad X \gg Z_0$$

This corresponds to the case of a transmission line terminated in an impedance composed primarily of a large negative reactance. Therefore, such an impedance translated back to the apexes of the cones will give the input impedance of the antenna. Since we are considering the case of $r \ll \lambda/4$, the change of impedance will also be small and the input impedance will consist of a small resistance and a large negative reactance.

In the case when r equals $\lambda/4$ it is more difficult to judge what the reflection of the *TEM* wave should be because the ratios of ϵ_θ to \mathcal{C}_ϕ for the various free-space waves are not all the same. For the waves of the three lowest orders, approximate values¹ are

$$\begin{aligned} \frac{\epsilon_\theta}{\mathcal{C}_\phi} &= \sqrt{\frac{\mu}{\epsilon}} (0.7 - j.2) \\ \frac{\epsilon_\theta}{\mathcal{C}_\phi} &= \sqrt{\frac{\mu}{\epsilon}} (0.25 - j.8) \\ \frac{\epsilon_\theta}{\mathcal{C}_\phi} &= \sqrt{\frac{\mu}{\epsilon}} (0.1 - j1.5) \end{aligned}$$

For still higher order modes, the real quantity diminishes, and the reactive term increases in magnitude. Therefore, if the two lowest order waves predominate, the equivalent termination for the *TEM* wave has a large real component. However, it is more reasonable to suppose that the lowest order modes do not predominate, because the intensity of the field components of the *TEM* mode varies as $1/\sin \theta$, whereas the intensity of the corresponding field components of the lowest order free-space *TM* wave varies as $\sin \theta$, and the two do not correspond at all except in the region where $\theta = 90$ deg. Accordingly, the lowest order waves should dominate the summation for only extremely large-angle cones, if at all. Therefore, let us assume that waves other than those of the lowest order are important. The equivalent termination for the *TEM* wave is of the form $R - jX$ where R is small and X is large compared with Z_0 . Translating an impedance of this kind back to the input of the antenna to obtain the input impedance yields a small resistive component and a small positive reactive component since the length of the transmission line represented by the antenna is 90 deg.

¹ Computed from curves by Schelkunoff, *op. cit.*, Chap. XI, Sec. 2.

Recapitulating the conclusions drawn thus far we have the following:

1. For a cone of long side length, the input impedance approaches the characteristic impedance of the cone.
2. For a cone a quarter-wave long, the input impedance should be approximately real, with a resistive component less than Z_0 of the cone and a small positive reactive component.
3. For a very short cone, the input impedance should consist primarily of a large negative reactance.

It should be pointed out that the first conclusion is not valid for cones of very small angle. The term $1/\sin \theta$ (in the expressions for \mathcal{R}_0 and \mathcal{E}_0

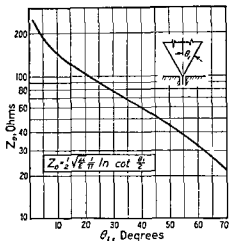


FIG. 4-7.—Computed characteristic impedance of transmission line consisting of a single conical conductor and an infinite ground plane

of the *TEM* wave) increases very rapidly as θ approaches 0 or π . Therefore, the boundary conditions can be met only by a summation in which very high-order free-space *TM* waves figure prominently, and these very high-order modes do not have radial impedances that approximate those of the *TEM* wave except at extremely large radii. All three conclusions are consistent with experimental data (presented in Sec. 4-3) for cones whose angles are between 5 and 45 deg, and they may be expected to apply also to larger angle cones.

These conclusions are useful in the design of cone antennas. One can choose a cone angle with the knowledge that as the frequency rises, the input impedance of the antenna will approach the value of the characteristic impedance of the cone transmission line as determined by Eq. (4-1). (Figure 4-7 shows a curve computed from this equation.) In fact, for cones of greater than 5-deg angle, the input impedance approaches closely the characteristic impedance as the length of the cone side becomes of the order of one wavelength. (This is consistent with the fact that for the first several free-space *TM* waves, the radial impedance approaches very close to the radial impedance of the *TEM* wave at these radii.) Two other generalizations can be drawn from the experimental data presented in Sec. 4-3.

1. The first resonance occurs when the cone is slightly less than $\lambda/4$ in length and the resistance is then somewhat less than the resistance of a thin stub at first resonance.

2. The second resonance occurs when the cone is slightly less than a half-wave long, and the lower the Z_0 , the lower the input resistance.

The foregoing discussion leads directly to the conclusion that a cone antenna that is used with a transmission line whose characteristic impedance is comparable with that of the cone, can readily be made to possess broad-band impedance characteristics.

An analysis of the performance of *cylindrical* antennas is considerably more difficult, because they do not possess characteristic impedances which remain constant throughout their length. Instead the characteristic impedance gradually increases with the distance from the feed point. Nevertheless, in general, the effect of increasing thickness is quite similar to the effect of increasing cone angle. Accordingly the input impedance of a thick cylinder changes less with frequency than does that of a thin one. The experimental data given in Sec. 4-4 show the behavior.

4-2. General Considerations: Patterns.—A general discussion of the effect of antenna thickness upon radiation patterns is given in Chap. 1, but a few remarks relating the preceding section on cones to pattern phenomena are of interest here. It will be recalled from Chap. 1 that the radiation pattern depends, among other things, on the magnitude of the reflected current relative to the transmitted current on the antenna, or in more general terms upon the phase and amplitude distribution of the current. It will also be recalled that, in the case of the cone, the reflected *TEM* wave becomes very small when the antenna exceeds a particular length measured in wavelengths. Therefore, the pattern might be expected to have a very pronounced tendency to point along the axis of the cone as it does in the case of a nonresonant wire. However, the situation is modified because the current on the antenna due to the *TEM* wave is not the total current. Actually, the *TM* waves within the boundary sphere are associated with currents flowing on the antenna. The reason these were not involved in the cone-impedance discussion was because they disappear at the cone input and, therefore, do not affect the input impedance. But since they are present on the antenna, they do affect patterns. This is mentioned here primarily because the discussion of the behavior of the *TEM* wave might otherwise erroneously suggest that the patterns could readily be computed from the behavior of the *TEM* wave. An accurate calculation of the patterns requires that the distribution of the free-space *TM* modes be determined in accordance with boundary requirements.

The effect of bending a dipole antenna to form a V is well known. But patterns of a V composed of thick elements are quite different at high frequencies (because of the guiding effect) from those of a thin-wire V.

The subject is treated from an experimental viewpoint in the following sections.

4-3. Cone Antennas.—In a general way the performance of cone antennas is described in the preceding sections. The present section is devoted primarily to a discussion of experimental data and the description of some typical practical antennas of this type.

In a study made at the Radio Research Laboratory, the input impedance of a number of single-cone antennas was measured. The curves shown in Figs. 4-8 to 4-10 give the results obtained. The first two of

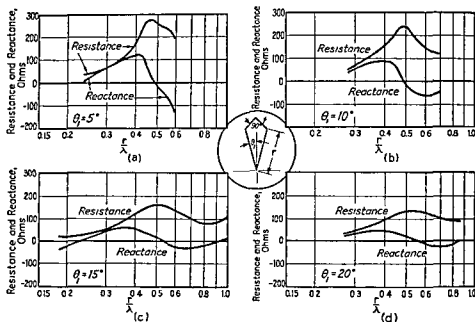


FIG. 4-8 — Measured resistance and reactance (vs. length of cone side) of conical antennas showing effect of varying cone angle.

these figures show the impedance of a family of cones as the cone angle varies from 5 to 47 deg. The caps for these did not consist of spherical sections, but were made of inverted cones with a total included angle of 90 deg (see sketches accompanying the curves). Figure 4-10 shows the effect of different caps upon the input impedance of a cone for which $\theta_1 = 30^\circ$. Four cases are presented: that of no cap, a 90-deg conical cap, a hemispherical cap (center in plane of widest part of cone), and a spherical cap that coincides with the boundary sphere.

Since the 90-deg conical cap causes the total length of the antenna to be somewhat greater than it would be if the cap were coincident with the sphere that terminates the input cone, it is clear that the data of Figs. 4-8 and 4-9 cannot be exactly applied to the cones discussed in Sec. 4-1. However, the physical difference between the two is fairly small when

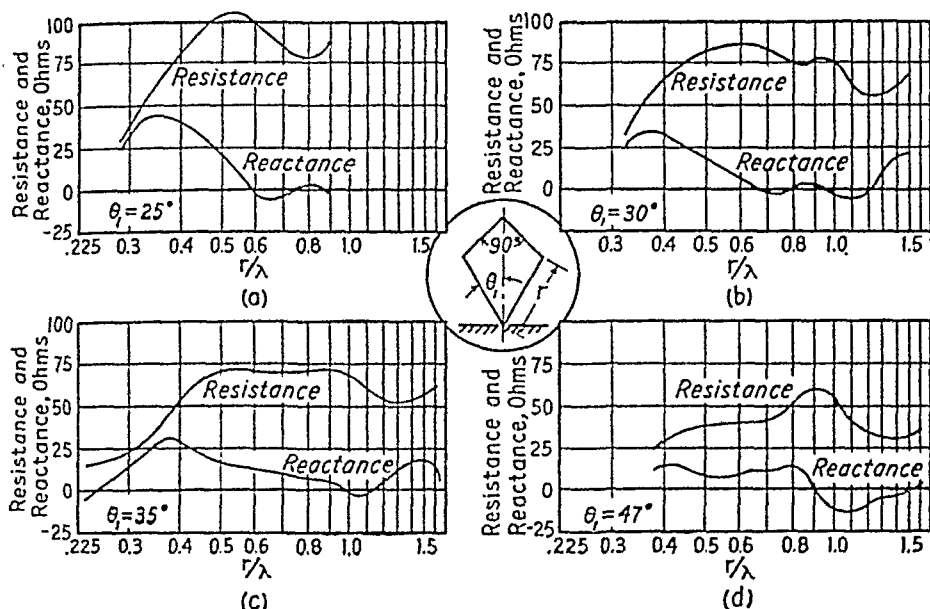


FIG. 4-9.—Measured resistance and reactance (vs. length of cone side) of conical antennas showing effect of varying cone angle.

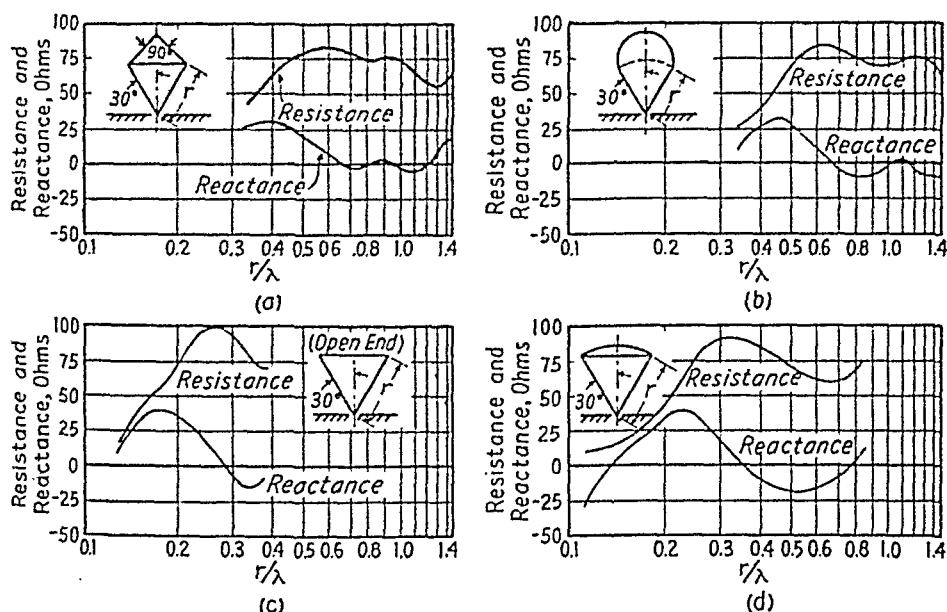


FIG. 4-10.—Measured resistance and reactance (vs. length of cone side) of cone antennas showing effect of various caps.

the cone angle is small, and the results can be compared. It will be noticed that the general trends mentioned in Sec. 4-1 are followed.

One of the most interesting quantities to consider in connection with these data is the magnitude of the reflection coefficient for the *TEM* wave on the cones themselves. This can readily be computed from the input impedance and the characteristic impedance of the cones. Figure 4-11 shows the results of such a computation for the cones with the 90-deg caps. Two very definite trends are apparent. The first of these is that

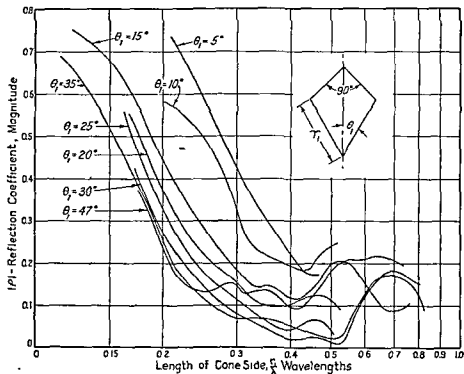


FIG. 4-11.—Magnitude of reflection coefficient for *TEM* wave on conical antennas of various angles vs length of cone side in wavelengths. Computed from data of Figs 4-8 and 4-9.

for all the cones the reflection coefficient starts at a large value and very rapidly decreases as the cone length increases. The second is that the reflection coefficient decreases as the cone angle increases. It will also be noticed that the reflection coefficient does not decrease continuously with increase of cone length, but instead fluctuates between zero and about 0.2 once it becomes small. These fluctuations may be due, at least in part, to experimental error. Possibly, however, the fact that the caps are not spherical is responsible for the fluctuations.

Another interesting quantity to consider is the variation of the input impedance at second resonance, *i.e.*, half-wave resonance, with the varia-

tion of the cone characteristic impedance. This has been plotted in Fig. 4-12. This curve for the region in which the data were taken may be closely approximated by the expression

$$Z_{\text{input}} = 1.5Z_0 - 30 \text{ ohms}$$

From Fig. 4-10, which shows the effect of different caps, one obvious fact is of interest. The general nature of the four curves is similar; the principal difference is that the antennas with the larger caps have the properties of longer antennas.

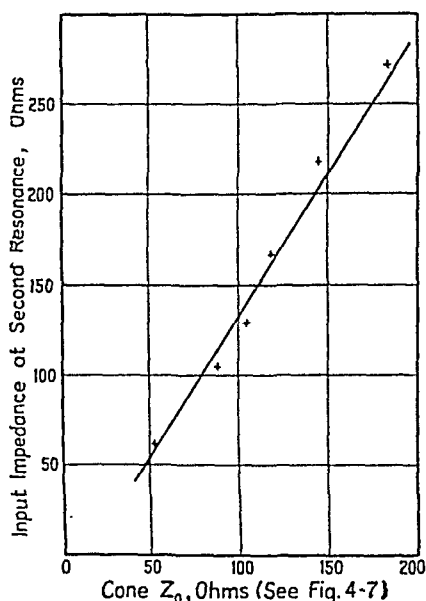


FIG. 4-12.—Measured input impedance at second resonance of single conical antenna vs. characteristic impedance (for *TEM* wave on cone).

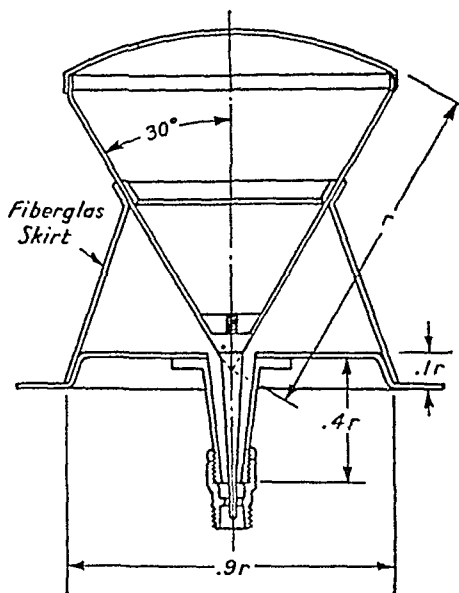


FIG. 4-13.—Cross-sectional view of conical antenna.

The information presented thus far is sufficient to permit the design of broad-band conical antennas as far as input impedance is concerned. To illustrate typical treatments of the practical problems presented, three specific antennas are discussed: the first of these is a single cone designed to operate against a ground plane to which it is normal, the second is a biconical dipole, and the third is a single cone like the first except that it is mounted at an oblique angle with the ground plane.

Figure 4-13 shows a cross-sectional view of a single-cone antenna with a spherical cap of the type discussed in the theory. This antenna is designed to match a 50-ohm line. Accordingly, Fig. 4-14 shows the voltage standing-wave ratio measured on a 50-ohm line terminated by the antenna. It will be noticed that the standing-wave ratio is less than 2.2:1 over the entire 10-to-1 frequency range measured. There are

several features of the design that warrant discussion. First, the angle of the cone is 30 deg, as shown in Fig. 4-13, even though a cone of this angle has a characteristic impedance of about 78 ohms (Fig. 4-7), rather than the 50 ohms that one might expect to have been chosen to match a 50-ohm line. This cone angle was chosen because the resistance of a 50-ohm cone falls below 50 ohms in the quarter-wave region more rapidly than does that of the 78-ohm cone. Therefore, a cone of characteristic impedance greater than 50 ohms will afford a more satisfactory match to a 50-ohm line at lower frequencies (for a given size antenna) than will a 50-ohm one. However, at higher frequencies, the higher impedance cone will not be so satisfactory. In fact, the voltage standing-wave ratio for the 78-ohm cone may be expected to approach $\sqrt{3}/2$ or approximately 1.6 as the frequency rises.

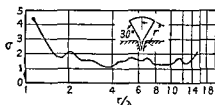


FIG. 4-14—Measured voltage standing-wave ratio on 50-ohm coaxial transmission line terminated by the antenna of Fig. 4-13

Nevertheless, if 1.6 is an acceptable standing-wave ratio, the 78-ohm cone represents a better design compromise, since it results in a smaller antenna for a given low-frequency limit. In the case of this particular antenna, an additional feature is present. This consists of a tapered section of coaxial line in series with the input to the cone. The impedance of this section varies from 50 ohms at the input end to approximately 78 ohms at the antenna end. Its length has been so chosen that at low frequencies it is too short to have much effect, while at high frequencies it transforms the 78 ohms of the cone down to 50 ohms at the input. Thus the advantage at low frequencies of the higher impedance cone is maintained, but the disadvantage at the high-frequency end is compensated.

Another feature of interest is the mechanical means for supporting the antenna. This consists of a thin Fiberglass skirt, sufficiently removed from the base of the antenna to be in a comparatively weak field. Because of its thinness and its location, the reflection that it produces is small. Since it provides adequate mechanical support, it removes the necessity for filling the tapered section with dielectric material. If the taper were filled with solid dielectric, there would be a source of reflection at the base of the cone due to the junction between solid dielectric and air. As it is, there is a similar discontinuity due to the boundary between the air in the taper and the solid dielectric material in the transmission line leading to the antenna. But the reflections due to this discontinuity are less severe than they would be if the discontinuity were at the other end of the taper, because the smaller the diameter of the section at which a junction of this type occurs, the smaller the reflections it

introduces. Thus, although several discontinuities are present (including a step in the base plate that anchors the Fiberglass skirt), the design is such that they are fairly small, and reference to the standing-wave ratio curve shows that the broad-band possibilities of the cone are substantially realized.

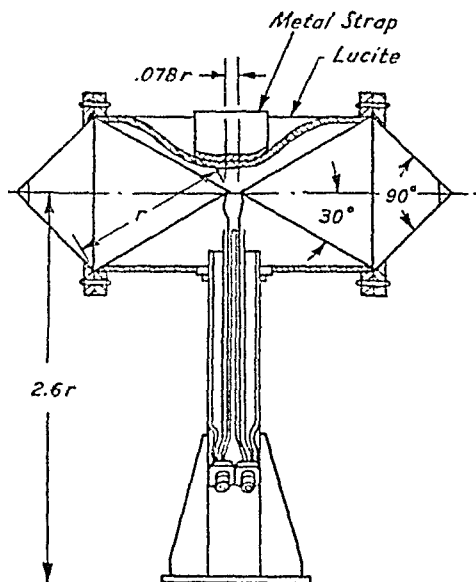


FIG. 4-15.—Cross-sectional view of biconical antenna.

Figure 4-15 shows a cross-sectional view of a biconical dipole. It will be seen that the antenna is straightforward, consisting of two conically capped cones, supported at their widest points by a Lucite cylinder and fed at the center by a 100-ohm line, consisting of two 50-ohm coaxial lines in parallel. A point of interest is that the metal strap, which encircles the plastic cylinder and was added for mechanical reasons, has only a minor effect on the performance. As far as input impedance is concerned, this may be judged from Fig. 4-16, which shows the voltage standing-wave ratio measured on a balanced 100-ohm feed line. The patterns for this antenna are not given here, since they are quite similar to the patterns given in Fig. 4-19 for a biconical antenna, without the strap, and with spherical caps.

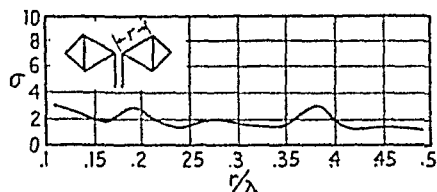


FIG. 4-16.—Measured voltage standing-wave ratio on balanced 100-ohm line terminated by antenna of Fig. 4-15.

Figure 4-17 shows a cone that has been tilted with respect to its mounting surface, in order to fill in the deep null that would otherwise exist in the radiation pattern in the direction of the cone axis (see Sec.

1-6). Aside from the difference in base plates, it is almost identical with the first cone described (Fig. 4-13). The only other important difference is that the taper transforms from about 63 ohms to 50 ohms instead of from 78 ohms to 50 ohms. This change was required because the input impedance of the cone is lowered by its greater proximity to the ground plane. Figure 4-18 shows the voltage standing-wave ratios for this antenna measured on

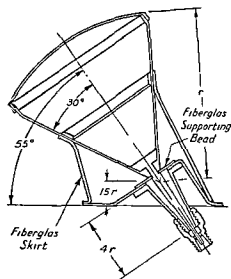


FIG. 4-17.—Cross-sectional view of tilted conical antenna

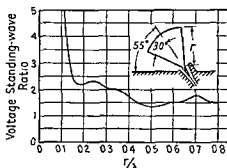


FIG. 4-18—Measured voltage standing-wave ratio on 50-ohm coaxial line terminated by antenna of Fig. 4-17.

a 50-ohm line. The patterns correspond closely to the patterns that follow for a biconical V antenna except, of course, that only half the pattern exists for a single cone against a ground plane.

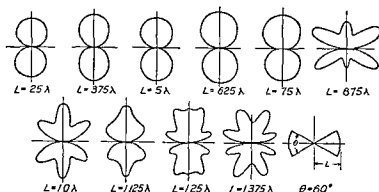


FIG. 4-19.—Measured radiation patterns of collinear biconical antenna.

Figures 4-19 and 4-20 show two sets of measured radiation patterns. In both cases the antenna consisted of a biconical radiator composed of two identical spherically capped cones. These caps coincided with the boundary sphere, and the cone angle, measured from the cone axis to the

cone surface, was 30 deg. For the first set the cones were collinearly mounted, and for the second set they were mounted with their axes intersecting at an angle of 90 deg. In both cases the patterns were measured in a plane containing the axes of the cones.

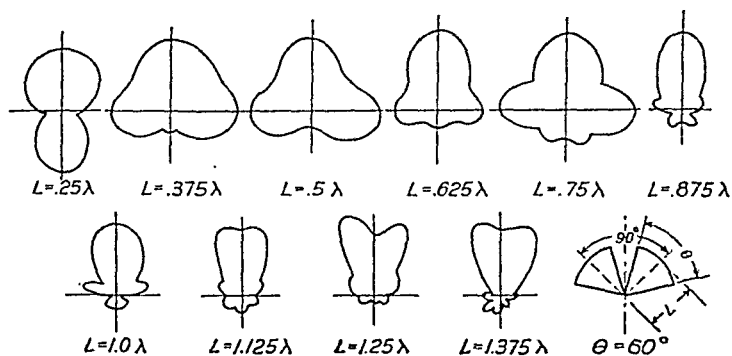


FIG. 4-20.—Measured radiation patterns of biconical V antenna.

For comparison, the patterns of a thin-wire linear dipole are shown in Fig. 4-21. These were computed by means of the expression¹

$$\text{Relative field} = \frac{\cos \frac{2\pi L}{\lambda} - \cos \left(\frac{2\pi L}{\lambda} \cos \theta \right)}{\sin \theta}$$

where L is the length of each half of the dipole and θ the angle from the antenna axis.

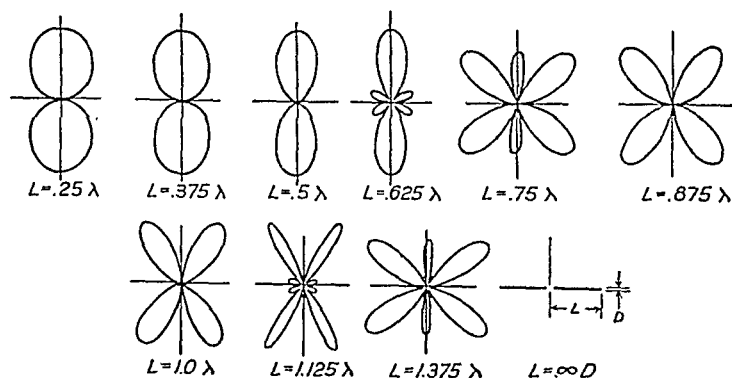


FIG. 4-21.—Computed radiation patterns of infinitely thin collinear balanced radiator.

In Fig. 4-19 it will be noticed that the patterns of the biconical linear antennas differ substantially from those of the thin dipole, except when the length of each half is of the order of $\frac{3}{8}$ wavelength or less. It will be noticed that the cone radiation becomes multilobed as the length of

¹ See, for example, Terman, F. E., "Radio Engineers' Handbook," Sec. 11, Par. 11, McGraw-Hill Book Company, Inc., New York, 1943.

the antenna increases, as does that of the dipole. However, neither the shape nor the number of lobes is in correspondence with those of the thin dipole, and furthermore, the depth of the nulls between lobes is much less in the case of the conical antenna. In Fig. 4-20, the most striking feature is the extent to which the radiation is concentrated in the direction toward which the cones are bent. The patterns for the longer cones resemble

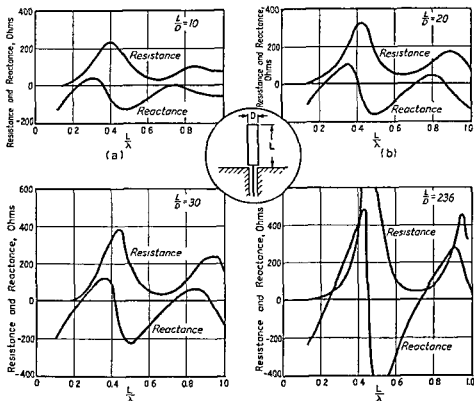


FIG. 4-22.—Measured resistance and reactance of cylindrical radiators vs cylinder length, showing effect of antenna diameter

horn patterns and indicate that the waves are guided along between the facing surfaces of the cones.

4-4. Cylinders.—In the preceding section cones were discussed. This section is concerned with data for cylindrical antennas and descriptions of a few typical antennas.

The Radio Research Laboratory has conducted an experimental investigation of the effect of thickness upon the input impedance of cylindrical radiators. Figure 4-22 shows the values obtained for a number of single cylinders fed against a ground plane. With these, as with the cones, the effect of increasing thickness is to smooth the curves,

primarily by lowering the magnitude of the resistance and reactance at the resonant peaks.

The curves reveal several interesting trends, which can be seen clearly in Figs. 4-23 to 4-25 in which some of the data have been replotted in different form. Figure 4-23 shows resistance at resonance vs. the ratio of length to diameter. It reveals that at first resonance the resistance varies only slightly with the thickness and is in the region of 35 ohms. At second resonance, *i.e.*, in the region where the length of the antenna is a half-wave, the resistance increases rapidly as the antenna becomes thinner. The behavior at third resonance is much like that at first

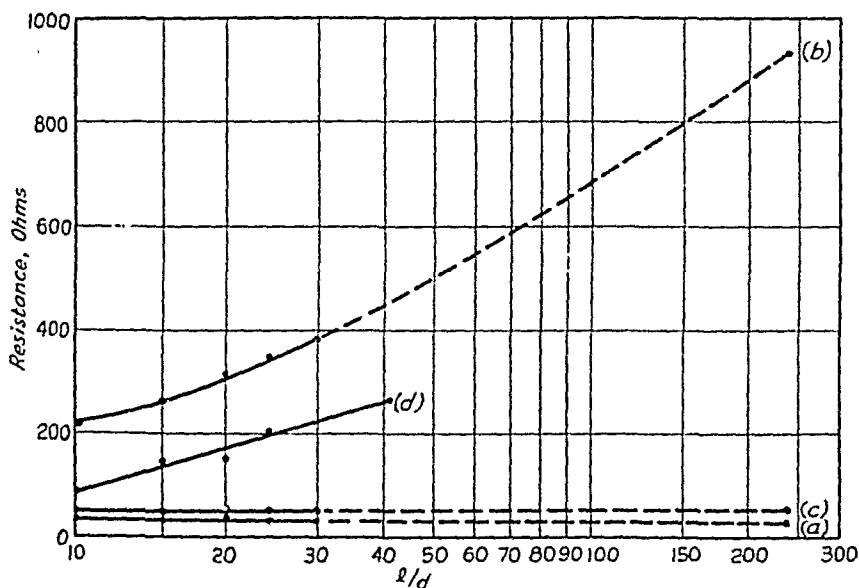


FIG. 4-23.—Measured resistance at first four resonances vs. the ratio of length to diameter for cylindrical radiators: (a) quarter-wave; (b) half-wave; (c) three-quarter-wave; (d) full-wave-resonance.

resonance, except that the average value is closer to 50 ohms. At fourth resonance, there is again considerable change with thickness, and over the range measured the resistance increases as the antenna becomes thinner.

Figure 4-24 shows the value of resistance at the first two maximums and at the first minimum vs. the ratio of length to diameter. These curves are very much like curves *b*, *c*, and *d* of Fig. 4-23, since the maximums and minimums occur quite close to the resonant points. However, since they do not occur exactly at the resonant points, the curves of Fig. 4-25 have been plotted to show the variation of the physical length of the antenna for various resonance effects as a function of thickness. In general, it is clear from these curves that the thicker the cylinder, the greater the effective electrical length for a given physical length.

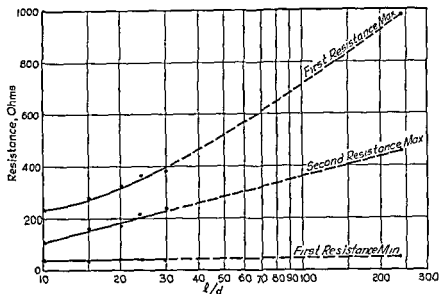


FIG. 4-24.—Measured resistance at first two maximums and at first minimum vs ratio of length to diameter for cylindrical radiators

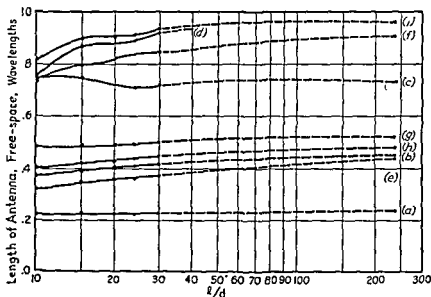


FIG. 4-25.—Physical length of cylindrical antenna for various resonance effects vs ratio of length to diameter: (a) quarter-wave; (b) half-wave; (c) three-quarter-wave; (d) full-wave resonance; (e) first; (f) second positive reactance peak; (g) first negative reactance peak; (h) first; (i) second resistance peak.

In all cases investigated, except that of the thinnest radiator, the diameter of the antenna was greater than the diameter of the center conductor of the coaxial line with which the antenna was energized. Clearly the capacitance between the ground plane and the base of the antenna

cylinder cannot be neglected. The procedure employed in obtaining the data consisted of measuring the approximate admittance due to this base capacitance and subtracting it from the total measured admittance. This value of admittance was obtained by replacing the antenna with a very thin metal disk of the same diameter as the antenna base and at the same distance from the ground plane and measuring its admittance.

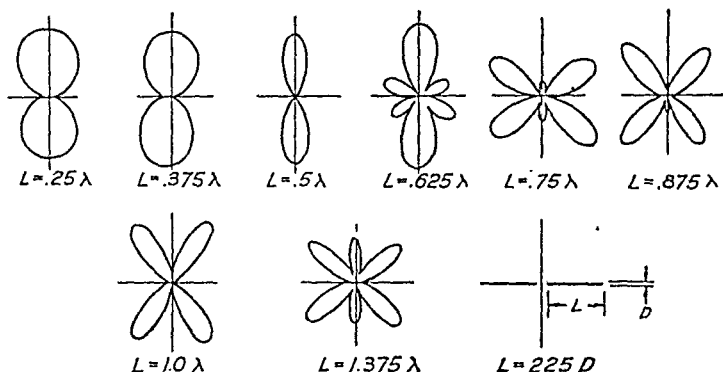


FIG. 4-26.—Measured radiation patterns of collinear cylindrical balanced radiator. Length-to-diameter ratio (each half) equal to 225.

It is evident that the value thus obtained is somewhat too large, since it included the admittance between the upper side of the disk and the ground screen, but the error involved is probably small and, in any case, would not seriously affect the general trends. In each case the cylinders were capped by a flat disk, although the presence of a cap made very little difference.

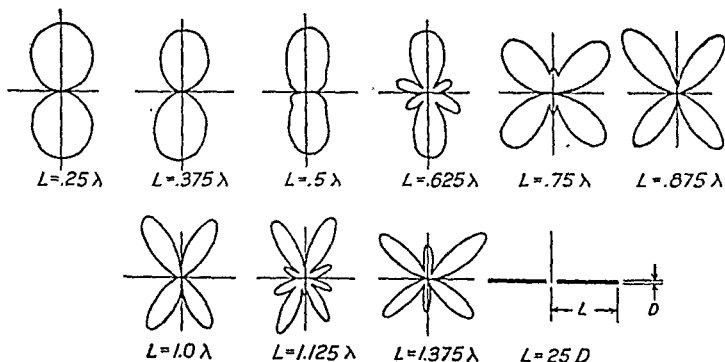


FIG. 4-27.—Measured radiation patterns of collinear cylindrical balanced radiator. Length-to-diameter ratio (each half) equal to 25.

Figures 4-26 to 4-29 show a set of measured patterns for a series of linear dipoles of varying thickness. A comparison between these and the computed patterns of Fig. 4-21 for a thin dipole shows that when the antenna is short the radiation pattern is virtually independent of thickness, but that as the antenna length increases above about a half-wave

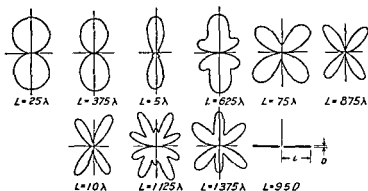


FIG. 4-28—Measured radiation patterns of collinear cylindrical balanced radiator. Length-to-diameter ratio (each half) equal to 9.5.

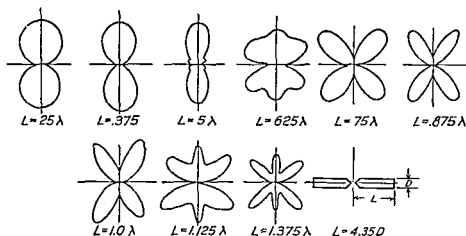


FIG. 4-29—Measured radiation patterns of collinear cylindrical balanced radiator. Length-to-diameter ratio (each half) equal to 4.35.

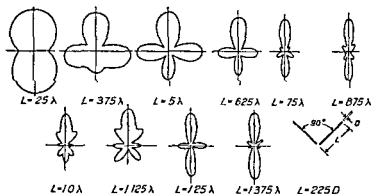


FIG. 4-30.—Measured radiation patterns of cylindrical V dipole. Length-to-diameter ratio (each half) equal to 225.

(for each half) the effect of thickness variation is marked. In general, increased thickness of the antenna is accompanied by a "filling in" of the nulls between lobes. Figures 4-30 to 4-33 show a set of measured

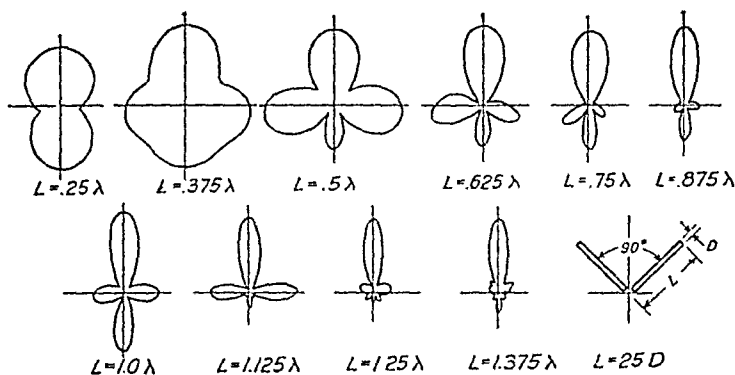


FIG. 4-31.—Measured radiation patterns of cylindrical V dipole. Length-to-diameter ratio (each half) equal to 25.

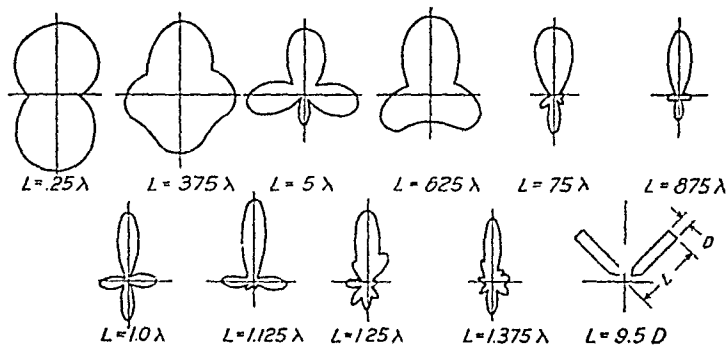


FIG. 4-32.—Measured radiation patterns of cylindrical V dipole. Length-to-diameter ratio (each half) equal to 9.5.

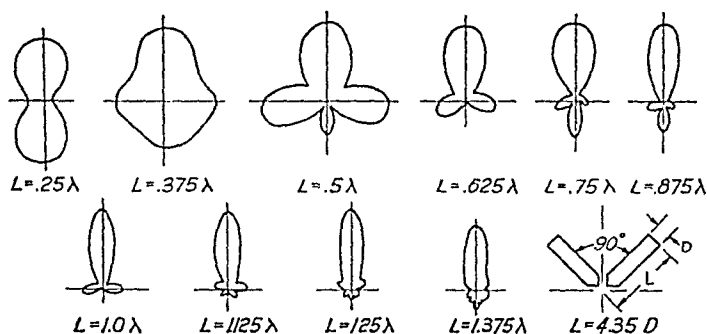


FIG. 4-33.—Measured radiation patterns of cylindrical V dipole. Length-to-diameter ratio (each half) equal to 4.35.

patterns for a series of 90-deg V antennas, composed of the same elements as were used for the linear dipoles mentioned above. The patterns for the thicker elements are quite similar to those of Fig. 4-22 for the biconi-

cal V, i.e., with increasing frequency the radiation becomes concentrated in the direction faced by the aperture between the two antenna elements. However, for the V composed of the thinnest elements, in general, no such effect is evident.

The preceding information is sufficient to permit a comparison between the cone and cylinder as broad-band radiators. Such a comparison leads to the conclusion that the cylindrical shape is not so readily adapted to broad-band practice as is the cone since its impedance is not so constant with frequency. In fact the cylinder is rarely used when maximum band width is important except when either (1) the greater bulk of the cone (for a given low-frequency limit) precludes its use, or (2) the pattern characteristics of the cone restrict

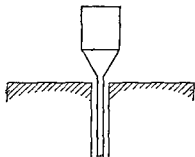


FIG. 4-34.—Thick cylindrical radiator with conical base.

its usefulness, in the particular application, to that band width which can be achieved with the cylinder. A shape that is frequently more useful than the cylinder, since its performance approximates that of a cone, while its bulk is not so great, is the combination of thick cylinder and conical base, shown in Fig. 4-34. In many cases this same shape is

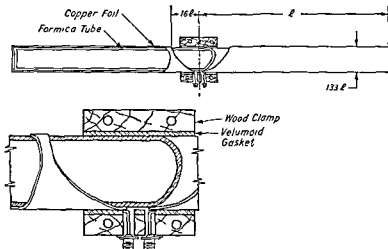


FIG. 4-35.—Cylindrical dipole.

arrived at for a different reason: to eliminate the large lumped capacitance to ground that would be present if the cylinder were not tapered.

Nevertheless, simple fairly thick cylinders are frequently used as moderately broad-band radiators. The band width of such antennas can usually be improved by means of suitable compensating networks. The

two examples that follow illustrate the use of simple uncompensated radiators.

Figure 4-35 shows a cylindrical center-fed dipole of total length equal to 16.5 times the diameter. The curve of Fig. 4-36 shows the measured voltage standing-wave ratio. This was not measured on the 100-ohm balanced line, but rather on a 50-ohm coaxial line, at the input to a type-II balun used to convert from unbalanced line to balanced line as shown. Because of the presence of the balun, the standing-wave ratio is not actually that of the antenna, although the order of magnitude is the same, and the curve, therefore, gives a good idea of the quality of the match achieved.

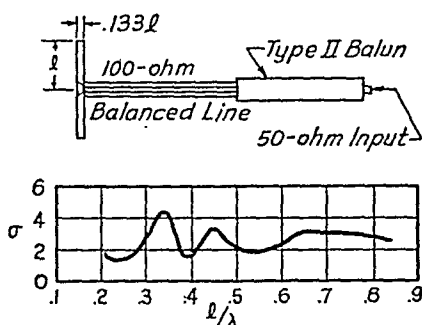


FIG. 4-36.—Measured standing-wave ratio of antenna of Fig. 4-35.

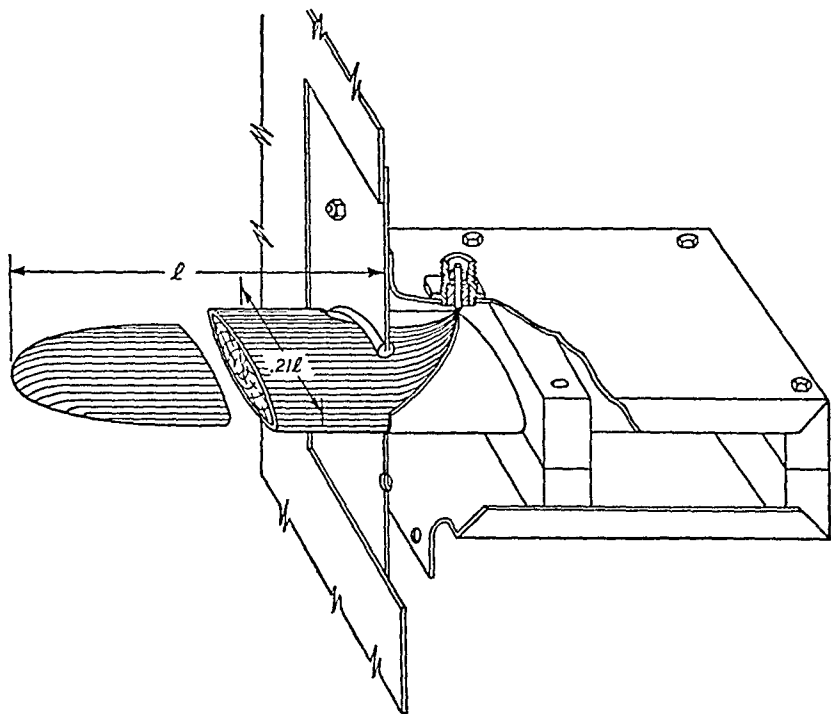


FIG. 4-37.—Streamlined stub antenna consisting of copper-covered wooden mast.

Figure 4-37 shows a thick stub designed to work against a ground plane. In this case the shape is not circularly cylindrical¹ since the

¹ No quantitative data are available that show the effect of a change of this type in the shape of the cross section of a cylindrical radiator. However, the effect of such

stub is streamlined for use on aircraft. Physically, the antenna consists of a piece of copper-plated plastic-impregnated wood. One of the interesting features is the way in which the connection is made between the base of the stub and the coaxial line. Reference to the figure shows that the triangular strip of metal and the metal clamp that supports the antenna constitute a crude tapered section of line. The voltage

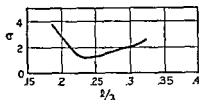
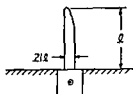


FIG. 4-38.—Voltage standing-wave ratio on 50-ohm line terminated by antenna of Fig. 4-37.

standing-wave ratio of this antenna is presented in Fig. 4-38. It will be noticed that a moderate band width is achieved despite the extent to which elegant electrical design has been sacrificed to achieve mechanical strength and simple construction.

4-5. Effect of Reflectors on Cones and Cylinders.—Quantitative data relating to the effect of reflectors upon cones and thick cylinders are not available. It is, therefore, necessary, in current practice, to be guided by very general principles and analogy with the behavior of thin-wire radiators.

So far as the impedance is concerned, it is clear that another conductor in the field of the antenna represents a discontinuity in the medium through which the waves are propagated. From the analogy between an antenna and a transformer, it is apparent that a discontinuity of this type, like any other discontinuity, produces reflected waves that will appear at the input side of the transformer. Further, since the distance in wavelengths between this source of reflection and other sources of reflection in the transformer varies with frequency, the total effect will be to cause the input impedance of the antenna to show more variation with frequency than would be the case if the reflector were not present. Accordingly, the general effect of a reflector upon input impedance is to decrease the width of the band within which a given degree of match is obtainable.

a deviation is probably small, and the antenna should have properties that approximate those of a circular cylinder of which the diameter is equal to the average diameter of the streamlined cross section.

CHAPTER 5

SLEEVE ANTENNAS

By E. L. BOCK, J. A. NELSON, AND A. DORNE

5-1. General Considerations.—An electromagnetic radiator is termed a *sleeve antenna* when it incorporates a tubular conductor, *i.e.*, sleeve, of which the exterior is used as a radiating element and the interior as the outer conductor of the coaxial transmission line that feeds the

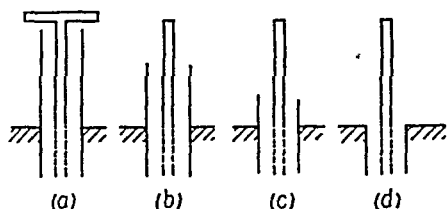


FIG. 5-1.—Sleeve-stub antennas with uniform cross section.

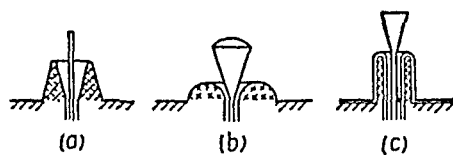


FIG. 5-2.—Sleeve-stub antennas with non-uniform cross section.

antenna. The length of the sleeve may be any portion of the total length of the antenna. This is illustrated by the four antennas of Fig. 5-1. Figure 5-1a shows the extreme in which the sleeve constitutes the entire radiating portion of the antenna. An antenna of this type is frequently termed *end fed*. Figure 5-1d shows the other extreme, in which the sleeve length has become zero and the sleeve antenna is reduced to an ordinary cylindrical stub.

An unlimited number of sleeve-antenna configurations is possible. A few examples are given in Fig. 5-2. It will be noticed that each of these has been shown to possess a sleeve with the inner diameter substantially less than the outer. In each case a crosshatched region between the walls of the sleeve is shown. It is evident, from consideration of the structures, that the performance of the antennas will in no way depend upon whether the crosshatched regions are empty or are filled with metal, since they are separated from all fields by continuous metal surfaces. This matter has been brought out to distinguish the sleeve antenna as defined here from the type of antenna shown in Fig. 5-3, which also involves a "sleeve" element but which differs in that the sleeve is used as a choke and the space corresponding to the crosshatched regions of



FIG. 5-3.
—Center-fed
coaxial di-
pole.

Fig. 5-2 must be empty in order that the choking effect take place. Antennas of this latter type are usually forms of center-fed cylindrical dipoles, and the purpose of the sleeve is merely to permit feeding them with coaxial transmission line.

It will be noticed that the sleeve antennas discussed thus far are all unbalanced in that the relation between antenna and ground plane corresponds to that which exists with a single cone or stub operated against a ground plane. Balanced-sleeve antennas are also widely used, and Fig. 5-4 shows several examples. These have been formed from the examples of Figs 5-1 and 5-2 by removing the ground plane in each case and replacing the image of the antenna by a real one whose configuration is the same as was that of the image. In the following discussion the unbalanced type will be referred to as *sleeve stubs* and the balanced type as *sleeve dipoles*.

The type of sleeve antenna most frequently used, and the type to which the following discussion is primarily devoted, is composed of elements of uniform cross section. Under some circumstances the use of tapered elements offers advantages from the point of view of impedance matching. This is particularly likely to be true when the diameter of the radiating elements is much greater than that of the feed line, since appropriate tapering avoids the large discontinuities that are otherwise almost inevitable. In many cases,

however, the advantages to be obtained by the use of tapered elements do not justify the complexities of construction that they entail, and consequently the nontapered sleeve antenna may be considered the basic type.

The antennas discussed below have in each case been made of the sleeve type because of the following two reasons:

1. The sleeve antenna permits unusually rugged mechanical construction with a minimum of conflict between electrical and mechanical design.
2. The sleeve antenna possesses impedance characteristics that, for broad-band purposes, compare favorably with those of other antennas of comparable bulk.

It will be noticed that these reasons do not include any reference to radiation patterns. This is because sleeve antennas that are short in

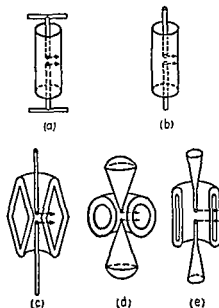


FIG. 5-4.—Sleeve-dipole antennas.

terms of wavelengths (as are those in these examples) provide patterns that are almost exactly the same as those of ordinary cylindrical antennas of the same lengths. Therefore, pattern considerations have not been involved in the choice of the sleeve design. In making a choice between long sleeve antennas and long cylindrical antennas, patterns might be an important factor because then the patterns do differ substantially.

5-2. Impedance and Pattern Characteristics of Cylindrical-sleeve Antennas.—Figure 5-5 shows a view of a sleeve-stub antenna whose elements consist of circular cylinders. It is evident that the size and shape of the radiating surfaces of the antenna depend upon four parameters: the over-all length of the antenna (L), the length of the sleeve (l), the diameter of the sleeve (D), and the diameter of the further section

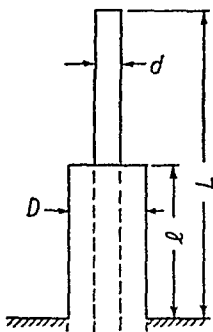


FIG. 5-5.—Cylindrical-sleeve-stub antenna.

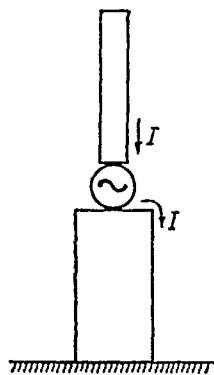


FIG. 5-6.—Schematic diagram of sleeve-stub antenna.

(d). These four quantities will determine the impedance and radiation-pattern characteristics of the antenna. The impedance referred to is the impedance that terminates the coaxial-line section enclosed by the sleeve. This is the input impedance of the radiating system proper. Thus, if a sleeve antenna were represented diagrammatically as in Fig. 5-6, the antenna impedance would be the impedance presented to the voltage source shown. This figure also serves to illustrate the fact that current flows on both sections of the antenna and that the sleeve is a device for displacing the feed point along the axis of a cylindrical (or conical, etc.) antenna.

Figure 5-7 shows the measured input resistance and reactance as defined above as functions of the over-all length of the antenna for two antennas like the one shown in Fig. 5-5. These antennas are identical except for the length l of the sleeve element, which is $\frac{1}{3}L$, and $\frac{2}{3}L$, respectively. For comparison, there is also shown (dashed lines) the resistance and reactance of the cylindrical antenna that would be obtained if the sleeve were removed, *i.e.*, when l equals zero. It will be

noticed that, for the region shown, the effects of increasing sleeve length include the following: (1) lowering the frequency at which the first resistance maximum occurs and, consequently, raising the resistance in the quarter-wave region; (2) producing an increasing capacitive reactance in the quarter-wave region.

These curves show the basic behavior of a sleeve antenna and make possible a comparison between it and simpler antennas such as the cone or cylinder. Since the curves were taken for particular values of D and d ,

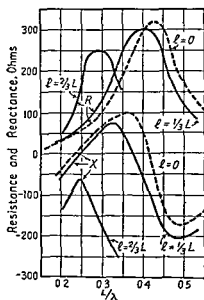


FIG. 5-7.—Measured input impedance of cylindrical-sleeve-stub antennas. $D = 0.15L$ (approximately) and $d = 0.406D$.

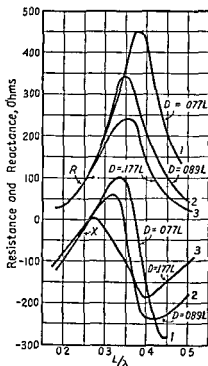


FIG. 5-8.—Measured input impedance of cylindrical-sleeve-stub antennas. $l = 0.5L$ and $d = 0.406D$.

it is obvious that they do not completely indicate the impedance trends of sleeve antennas. An idea of the effect of thickness variation can, however, be obtained from Fig. 5-8, which shows input resistance and reactance for three sleeve stubs of different thicknesses. It will be noticed that the effect of increasing thickness is similar to the corresponding effect in the case of simple cylindrical antennas. As the antenna becomes thicker, the maximum value of the resistance decreases and also the maximum and minimum values of the reactance become closer together.

We have stated that the input impedance characteristics of sleeve antennas are a function of four parameters. Practically, however, since

the section of transmission line enclosed by the sleeve is part of the antenna structure, the impedance at its input is likely to be of principal ultimate importance. Therefore, the characteristic impedance of this line is a fifth parameter which may be independently varied. The

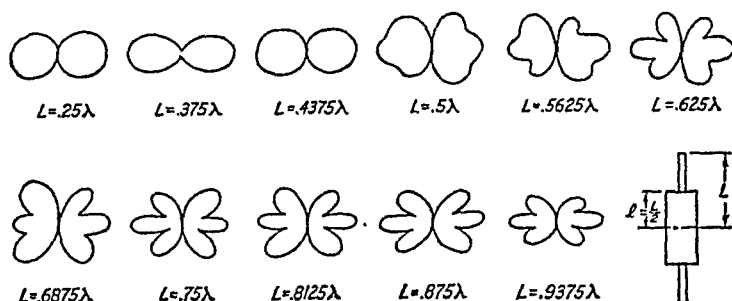


FIG. 5-9.—Measured field patterns of sleeve dipole. $l = 0.5L$, $D = 0.25L$, $d = 0.5D$.

existence of five parameters affords an unusual amount of control and largely accounts for the usefulness of sleeve antennas. In the following sections, a number of examples are presented, which illustrate typical results that may be achieved.

Figure 5-9 shows a number of measured radiation patterns for a sleeve dipole. If these patterns are compared with those of Fig. 4-21, for a thin cylindrical dipole, one particularly interesting difference will be observed, that the radiation pattern does not approach a null at right angles to the antenna axis anywhere in the range shown (from L equals $\frac{1}{4}\lambda$ to L equals $1\frac{5}{16}\lambda$), whereas it does approach a null in the case of cylindrical antennas as the length of each half approaches 1 wavelength.

5-3. Examples of Sleeve Stubs.—Figure 5-10 shows a cross-sectional view of a sleeve-stub antenna composed of circular cylindrical elements. It will be noticed from the figure that the internal structure of the antenna (shown by the solid lines) differs from that of the example of Fig. 5-5 in two respects: (1) the section of coaxial line enclosed by the sleeve is tapered, and (2) the upper section contains a short-circuited length of coaxial transmission line whose open end faces in the direction of the sleeve. The purpose of the tapering is to permit a gradual transition from the diameter of the sleeve to the much smaller diameter of the coaxial line used to feed the antenna. The short-circuited

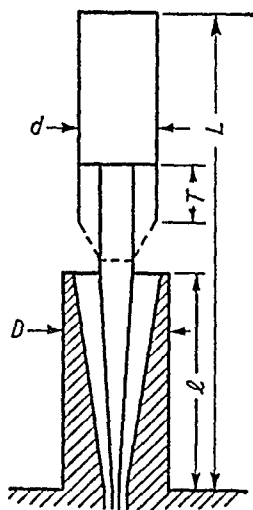


FIG. 5-10.—Cylindrical sleeve stub. $l = 0.46L$, $d = 0.162L$, $D = 0.216L$, $T = 0.123L$.

section of line in the upper portion of the antenna has as its purpose the introduction of an inductive reactance in series with the input impedance of the antenna. The desirability of this can be seen from the following discussion of the two curves of Fig. 5-11. The solid curve is the impedance,¹ plotted as a function of antenna length, at the input to the entire antenna assembly (shown by the solid lines in Fig. 5-10). The dashed curve gives the input impedance of the radiating structure alone, which was measured with the antenna modified in correspondence with the dashed line shown in Fig. 5-10 so that, for these measurements, the internal structure did not include the short-circuited section of coaxial line within the upper portion of the antenna. Also shown on the figure is a circle that encloses the region within which the impedance curve must fall if the voltage standing-wave ratio is to be less than 4:1 when the antenna is fed by a 50-ohm coaxial line. It will be noticed that the solid curve lies almost entirely within this circle, whereas the dashed one does not, thus showing that the impedance of the antenna assembly provides a better match to a 50-ohm line than does the impedance at the input to the radiating structure alone. Two impedance-compensating functions are performed that account for this difference. The first of these is performed by the short-circuited section of coaxial line, which, since it constitutes an inductance in series with the antenna input impedance, cancels some of the excessive capacitive reactance that can be seen to be present from the dashed curve. The second compensating function is performed by the tapered section of coaxial line, whose characteristic impedance has been made 80 ohms in order to complete the compensation partly affected by the series reactance.

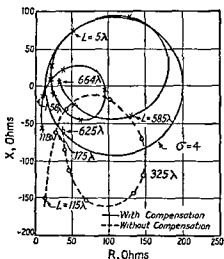


FIG. 5-11.—Measured impedance of sleeve-stub antenna of Fig. 5-10

constitutes an inductance in series with the antenna input impedance, cancels some of the excessive capacitive reactance that can be seen to be present from the dashed curve. The second compensating function is performed by the tapered section of coaxial line, whose characteristic impedance has been made 80 ohms in order to complete the compensation partly affected by the series reactance.

A somewhat different type of sleeve-stub antenna is shown by Fig. 5-12. In this example, as can be seen from its lateral cross section (Fig. 5-12b), the antenna is streamlined, and furthermore it is tilted at an

¹ In Fig. 5-11, the impedance is plotted on rectangular coordinates with the resistive component of the complex impedance as the abscissa and the reactive component as the ordinate. Thus, the length of the radius vector from the origin to any point is equal to the magnitude of the complex impedance. The graph obtained by plotting antenna impedance in this manner is often designated as an *R-X* diagram. *R-X* diagrams are discussed more fully in Chap. 2.

angle of 45 deg to the plane of its mounting base. The tilting results in a radiation pattern that is free from deep nulls through a 180-deg sector and, accordingly, permits more uniform coverage than would be available if the antenna axis were normal to its mounting base. Another point of difference between this antenna and that of the previous example is that the section of transmission line enclosed by the sleeve is of small uniform diameter, and consequently a large discontinuity does exist at the upper end of the sleeve. From the figure it can be seen that the plate that forms the upper end of the sleeve structure constitutes one of the two plates of a capacitor that is in parallel with the input (to the radiating surfaces). This capacitance has been deliberately introduced as one of the impedance-compensating elements of the antenna. The other compensating element is, as it was for the previous example, a series inductance located within the portion of the antenna farther from the base.

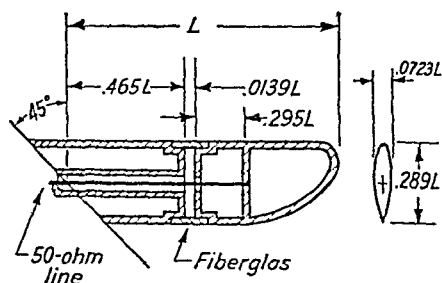


FIG. 5-12.—Inclined, streamlined, sleeve stub.

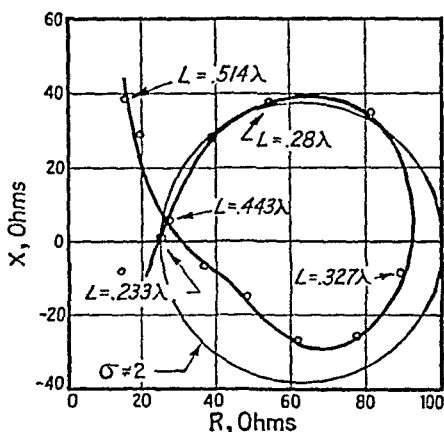


FIG. 5-13.—Measured input impedance of sleeve-stub antenna of Fig. 5-12.

Figure 5-13 shows the measured impedance of the antenna at the output end of the 50-ohm line enclosed by the sleeve. It can be seen that a portion of the impedance curve corresponding to a 1.9-to-1 frequency range falls within the 2:1 standing-wave-ratio circle for a 50-ohm line. By suitable adjustment of the parameters, the antenna could undoubtedly be designed to match lines of other characteristic impedances.

The fact that these two examples of sleeve stubs each includes a compensating series inductive element does not imply that elements of this type are always desirable. On the contrary, other forms of compensation are equally suitable, and in many cases adjustment of the five parameters originally mentioned may make special compensating elements unnecessary. In particular it should be noted that a series inductive element of the type described above precludes very large band widths, since the impedance of an antenna that includes such an element will approach an infinite value at the frequency at which the series element is $\frac{1}{4}$ wavelength long.

5-4. Simple Sleeve Dipoles.—Figure 5-4 shows a number of sleeve dipoles. Clearly, the input impedance of each of these assemblies consists of the sum of the individual impedances presented at the input to each of the coaxial lines that lead out through the two halves of the sleeve. Since these are equal, the input impedance of the dipole assembly consists of twice the input impedance to either section of sleeve-enclosed line. Accordingly, the input impedance of a sleeve-dipole assembly could be computed from curves, such as those of Figs. 5-7 and 5-8, by doubling the quantities obtained after the impedances shown by the curves had been translated through a section of coaxial line of the length and characteristic impedance of that enclosed by each half of the sleeve.

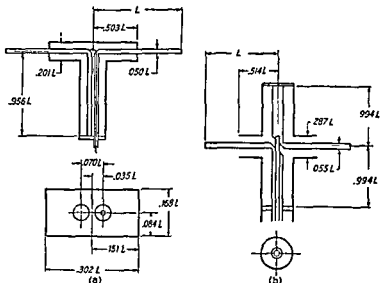


FIG. 5-14.—(a) Cylindrical-sleeve dipole with Type-III balun; (b) Cylindrical-sleeve dipole with Type-II balun.

If the antenna were fed by a balanced line, the value thus obtained would obviously constitute the terminating impedance for the line. Frequently, however, the antenna is fed with coaxial line, and the antenna structure is made to incorporate a balun. An antenna assembly of this type is shown by Fig. 5-14a. The dipole itself is composed of circular cylindrical elements, and the balun is type III in accordance with the classification of Chap. 3. It will be seen that the impedance of the short-circuited section of balanced line (of the balun) is in parallel with the impedance of the antenna proper. Obviously this parallel impedance must be taken into consideration in the design of the antenna. The curves of Figs. 5-15 and 5-16 show the relations involved and provide an example of one way in which the impedance of a sleeve dipole may be controlled. The first of these shows the impedance at the center of the

dipole for an antenna like that of Fig. 5-14a except that it does not include a balun. The curve of Fig. 5-16 shows the impedance for the complete antenna assembly, including the balun. A comparison of these two curves with the 50-ohm 2:1 standing-wave-ratio circle, which is also shown on each figure, indicates clearly that the balun exerts a substantial effect toward matching the antenna to a 50-ohm line. Obviously, if the balun were not present, the antenna proper would have to be of quite different proportions in order to effect as satisfactory a match.

In the preceding example, the sleeve consisted of a circular cylinder. There is, of course, no necessity for this shape, although it frequently leads to ease of fabrication. For example, the shape of the sleeve might

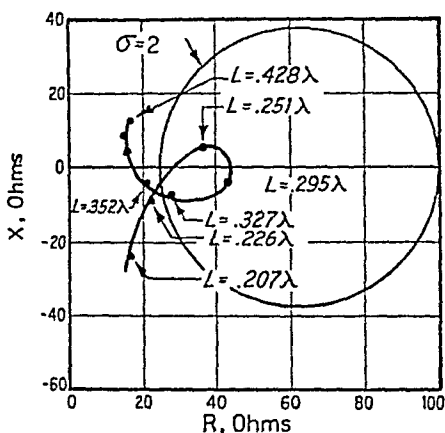


FIG. 5-15.—Impedance at center of cylindrical-sleeve dipole without balun.

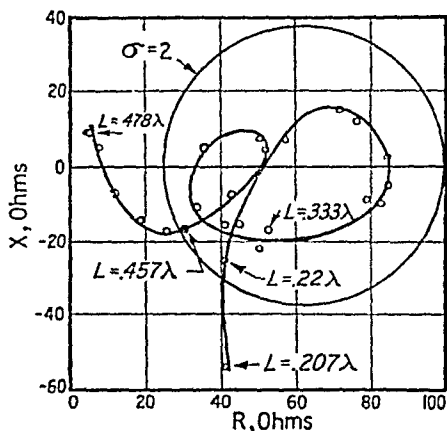


FIG. 5-16.—Measured impedance at center of cylindrical-sleeve dipole compensated for 50-ohm coaxial line.

be such that its transverse cross section were square. If all other design factors, including the characteristic impedance of the line within the sleeve, were unchanged, the effect upon the behavior of the antenna could be expected to be small. A comparison of this kind has been made experimentally, using the dimensions of the antenna of Fig. 5-14a, and the effect in that case was virtually negligible.

The discussion of the antenna shown by Fig. 5-14a has not thus far included any reference to a specific operating frequency. Rather it has been implied that the antenna may be made to function at any particular frequency by suitable adjustment of its size. Within rather broad limits this is correct, and if it were practicable to scale the antenna assembly completely, it may be presumed that the only practical limitation would occur when the antenna components became inconveniently small. Actually, however, when it is desired to increase the operating frequency, it is usually not feasible to scale the antenna completely, since the feed cable (including the length that extends through the balun) cannot be

arbitrarily chosen, but rather must exceed some minimum diameter in order either to meet power-handling requirements, or to be adequately rugged. In consequence, as the antenna is designed for higher and higher frequencies, the cross-sectional dimensions of the balun structure become proportionately larger. Experience indicates that the point at which the antenna will no longer function satisfactorily as a broad-band radiator

frequently falls at commonly used frequencies. Difficulty of this kind has been experienced, when the width of the balun shield was of the order of one-half the sleeve length. The difficulty may be circumvented to some extent by the use of a type-II balun. Figure 5-14b shows an antenna of this type. Its use doubled the center frequency of the band over which it was possible to attain satisfactory operation with a sleeve element of the proportions of that of Fig. 5-14a, when the feed cable was $\frac{3}{4}$ in. in diameter.

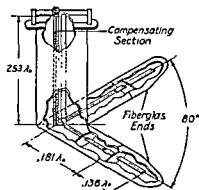


FIG. 5-17.—V-sleeve dipole.

5-5. Bent-sleeve Dipoles.—A sleeve dipole may be bent to form a V in order to decrease the directivity of its patterns. Figure 5-17 shows an antenna of this type designed for aircraft use and consequently streamlined. It is intended to be mounted on the underside of an aircraft,

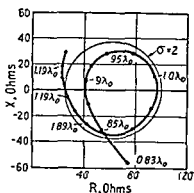
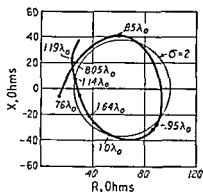


FIG. 5-18.—Measured impedance of V-sleeve dipole (a) before compensation; (b) after compensation.

and its radiation is then directed downward in a broad single-lobed pattern, which because of the V shape is free from deep nulls toward the sides. One of the aspects of the antenna which is of interest from a design standpoint is that it possesses a reasonable band width despite the small distance (0.21 wavelengths at the lowest frequency) between the dipole

element and the reflecting surface of the aircraft skin. The band width (for 50-ohm line) is about 1.38 to 1, as can be seen from the curve of Fig. 5-18a, which shows the input impedance of the antenna assembly as a function of frequency. A short series section of 77-ohm line has been included within the balun structure. The impedance-compensating effect which this introduces may be judged by comparing the curve of Fig. 5-18a (which shows the impedance at the center of the dipole, *i.e.*, before compensation) with that of Fig. 5-18b. It may be pointed out that, although the V shape as well as the proximity of the reflecting surface has tended to limit the band width of the antenna, there is no reason for presuming that substantially larger band widths are not obtainable with V-sleeve dipoles in free space.

The radiation patterns that are obtained with an antenna of this type may be judged from Fig. 5-19. Since these were measured with the antenna mounted on a small ground plane, they are not exactly the same as would be obtained with the antenna actually installed on an airplane. One of the principal differences is that the back lobes, which can be seen in Fig. 5-19b, are larger than those which would ordinarily exist in an actual installation. Nevertheless, the general shape of the patterns shown is the same as would be obtained in use.

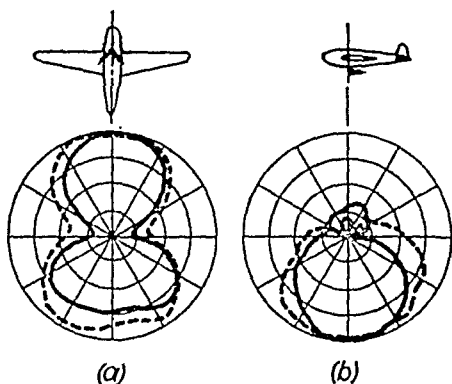


FIG. 5-19.—Radiation patterns of V-sleeve dipole (a) in plane of dipoles; (b) in plane bisecting the V. For dashed line $\lambda = 1.16\lambda_0$; for solid line, $\lambda = 0.84\lambda_0$. λ is the wavelength at which measurements were made; λ_0 the wavelength at approximately the center of the operating band.

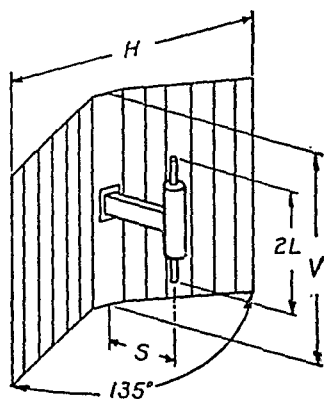


FIG. 5-20.—Sleeve dipole with corner reflector.

5-6. Sleeve Dipole-reflector Assemblies.

The sleeve dipole is frequently used in conjunction with a reflector in order to provide a directional radiator. Figures 5-20 and 5-21 show two broad-band assemblies of this type. The first of these utilizes a corner reflector, while the second utilizes simply a narrow grid of four rods. The radiation patterns obtained with antennas of these types are shown by Figs. 5-22 and 5-23. It will be noticed that the corner reflector provides patterns that are considerably more directional than are those of the other assembly. The principal advantage of the lower gain antenna is its smaller size. In

both cases, however, the radiation is essentially unidirectional throughout a 2-to-1 frequency range. In the case of the corner reflector, the pattern gain is quite uniform over the frequency band. Its maximum value (which occurs near the center frequency) is 8 db, and its minimum value (which occurs at the ends of the band) is 6 db, both quantities being referred to a half-wave dipole in free space.

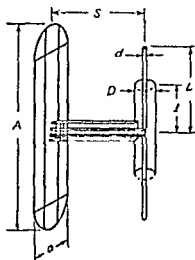


FIG. 5-21.—Sleeve dipole with narrow-grid reflector

In general, the standing-wave-ratio characteristics that can be obtained with a sleeve dipole used in conjunction with a reflector of either of the types shown are about the same as could be obtained with the same type of dipole without the reflector. The particular proportions required may be slightly different, but the band width obtainable

(for a given maximum standing-wave ratio) has in general proved to be about the same.

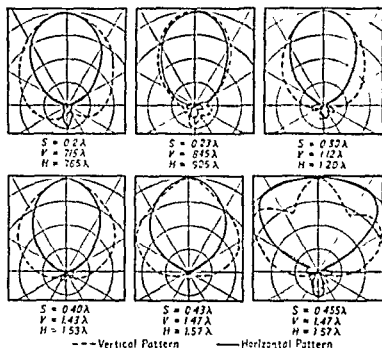


FIG. 5-22 —Field patterns of sleeve dipole in corner reflector.

The spacing between antenna and reflector is not particularly critical for either of these antenna assemblies. The lower limit to the spacing is imposed principally by the fact that, as spacing becomes very close, the resulting change in impedance cannot be compensated for by simple modification of the antenna. The upper limit to the spacing is imposed by pattern behavior. As the spacing approaches $\frac{1}{2}$ wavelength, the principal lobe begins to split. The dimensions of the reflector are not very critical. One of the principal requirements is that the dimensions

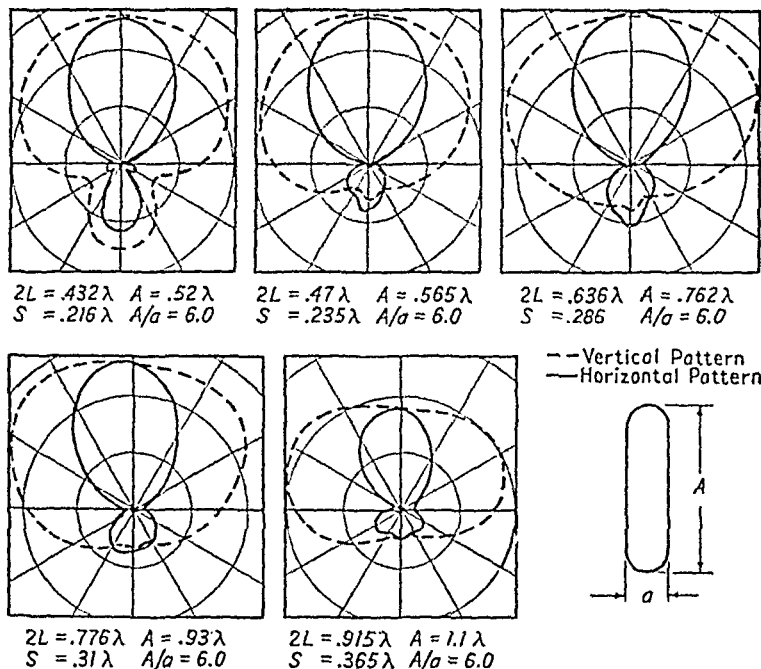


Fig. 5-23.—Field patterns of sleeve dipole and narrow-grid reflector.

parallel to the antenna axis must be somewhat greater than $\frac{1}{2}$ wavelength at the lowest frequency to be used.

5-7. Crossed-element Phased Antennas.—A crossed-element phased antenna of the type with which the following sections are concerned often consists of two dipoles crossed at right angles to each other (at the centers of the dipoles) and fed with currents equal in amplitude and 90 deg out of phase. If the two halves of each dipole are collinear, *i.e.*, the dipoles are straight, the antenna is frequently called a *turnstile*. If the plane of the turnstile is horizontal, then the horizontal radiation pattern of the antenna is nearly circular and the radiation in the horizontal plane is horizontally polarized. Along the axis of symmetry of the antenna, *i.e.*, directly above or below the antenna, the radiation is circularly polarized. Many other types of crossed-element phased antennas are possible. For example, two loops might be used instead of two dipoles. Then,

with the axis of symmetry of the antenna vertical as before, the radiation produced would be the same as that of the turnstile, except that the radiation in the horizontal plane would be vertically instead of horizontally polarized. Another possibility is to use two V dipoles, instead of

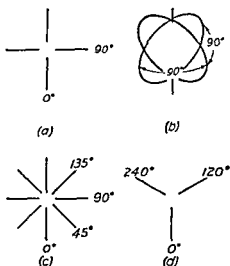


FIG. 5-24.—Types of crossed-element phased antennas relative phase angle of currents equal to angle between elements.

produces radiation patterns that are radially symmetrical. The polarization of the radiation directed along the axis of symmetry is circular, and the polarization of the radiation in other directions is elliptical or plane, depending upon the nature of the individual elements.

A symmetrical crossed-element phased antenna need not consist of two elements. A larger number may be used. The general requirements that must be met in order that the radiation pattern have an axis of symmetry along which the polarization is circular are as follows:

1. The angles between adjacent elements must be equal.
2. The currents in the elements must be of equal amplitude.
3. The phase difference between the currents on adjacent elements must equal the physical angle between the elements.

Figure 5-24 illustrates antennas of these types.

straight dipoles. If the planes of the individual V's intersect at right angles to each other along the center line of each V, then the antenna will have an axis of symmetry and the radiation pattern will be similar to that produced by the antennas of the preceding two examples, *i.e.*, the radiation will be circularly polarized in directions along the axis of symmetry and the radiation pattern will be nearly circular in the plane at right angles to the axis of symmetry. In this plane, however, the radiation will be elliptically polarized rather than plane-polarized as in the two previous examples.

In general, a radially symmetrical antenna of these or similar types

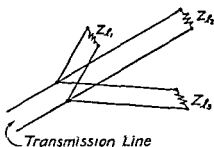


FIG. 5-25.—Diagrammatic representation of feed system of a crossed-element phased antenna, where Z_1 , Z_2 , and Z_3 represent the input impedances of the respective antenna elements.

Frequently, the individual elements of the antenna are fed by means of sections of transmission line connected in parallel at their input ends (see Fig. 5-25). Then the electrical length and impedance of these lines as well as the input impedance of the individual radiating elements control the amplitudes and phases of the currents on the individual elements. In the following section is discussed the problem of maintaining these amplitudes and phases constant over a broad frequency band. Crossed-element phased antennas are treated in this chapter principally because the sleeve dipole is one of the more suitable antennas of which they may be composed when broad-band performance is required.

5-8. Broad-band Crossed-element Phased Antennas.—The pattern characteristics of a crossed-element phased antenna may be maintained uniform over a broad band of frequencies through suitable control of the power distribution and the phasing between elements. The power-distribution requirement can be satisfied within reasonable limits by using broad-band dipoles that have small resistance variation over the frequency band.

Although the phasing is primarily dependent on the electrical length of the phasing line connecting the two elements, it can be shown that it is also dependent on the input impedance of the two parallel lines at the junction. Equation (5-10) is derived below for the antenna-element input impedance required to produce circularly polarized radiation when a given phasing line is used. The ideal antenna-element input impedance thus obtained is a function of the length and characteristic impedance of the phasing line.

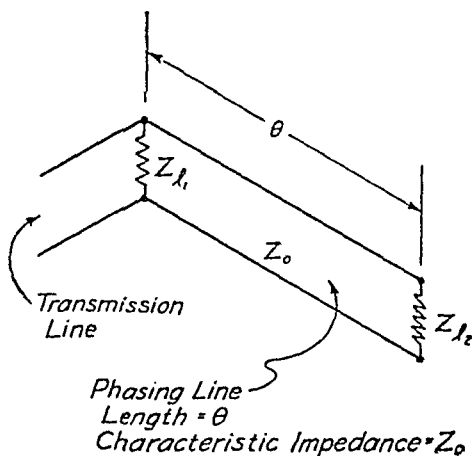


FIG. 5-26.—Diagrammatic representation of a two-element phased antenna.

Let us consider the conditions in Fig. 5-26. The input impedance of each of the two antenna elements is represented by Z_l since the two elements are identical. The elements are connected by a phasing line of characteristic impedance Z_0 and electrical length θ . The system is fed by a transmission line connected at the junction of the directly fed elements associated with Z_l and the phasing line. The reflection coefficient may be expressed as

$$\rho = \frac{B}{A} e^{j\psi} = \frac{Z_0 - Z_l}{Z_0 + Z_l} \quad (5-1)$$

where ρ is the complex reflection coefficient, B/A the ratio of the reflected wave to the direct wave on a transmission line of characteristic impedance Z_0 , and Ψ the phase angle of the reflection coefficient. If the current in the indirectly fed antenna is

$$i_{12} = A + Be^{j\Psi} \quad (5-2)$$

then the current and voltage on the phasing line at its input are

$$i_{\text{input}} = Ae^{j\theta} + Be^{j(\Psi-\theta)} \quad (5-3)$$

and

$$e_{\text{input}} = Z_0(Ae^{j\theta} - Be^{j(\Psi-\theta)}) \quad (5-4)$$

The voltage e across the phasing line is the same as that across the directly fed antenna; so the current in the directly fed antenna is

$$i_{11} = \frac{e_{\text{input}}}{Z_1} = Z_0 \left(\frac{Ae^{j\theta} - Be^{j(\Psi-\theta)}}{Z_1} \right) \quad (5-5)$$

Let

$$i_{11} = i_{12}e^{j\pi/2} \quad (5-6)$$

to fulfill the equal current and 90-deg phasing conditions. Then

$$\frac{Z_0}{Z_1} (Ae^{j\theta} - Be^{j(\Psi-\theta)}) = (A + Be^{j\Psi})e^{j\pi/2} \quad (5-7)$$

This may be simplified to

$$\frac{Z_0}{Z_1} \frac{e^{j\theta} - \frac{B}{A} e^{j\Psi} e^{-j\theta}}{1 + \frac{B}{A} e^{j\Psi}} = e^{j\pi/2} \quad (5-8)$$

Substituting the value of $(B/A)e^{j\Psi}$ from Eq. (5-1), we have

$$\frac{Z_0}{Z_1} \frac{e^{j\theta} - \left(\frac{Z_0 - Z_1}{Z_0 + Z_1} \right) e^{-j\theta}}{1 + \left(\frac{Z_0 - Z_1}{Z_0 + Z_1} \right)} = e^{j\pi/2} \quad (5-9)$$

Solving for Z_1 , we have

$$Z_1 = Z_0 \frac{\sin \theta - j \sin \theta \cos \theta}{\cos^2 \theta + 1} \quad (5-10)$$

Using the relationships given in Eq. (5-10), one may design an antenna with two crossed elements in space quadrature fed with equal currents in time quadrature, with good broad-band characteristics.¹

¹ It is, of course, possible to determine the ideal broad-band impedance curve to maintain current and phase relationships other than those of Eq. (5-10) by setting up Eq. (5-6) for the particular condition desired, substituting the values of i_{11} and i_{12} given in Eqs. (5-5) and (5-2) and solving for Z_1 .

The values of Z_i/Z_0 are plotted in Fig. 5-27 for values of θ from 45 to 135 deg and for relative frequency values from $0.5f_0$ to $1.5f_0$ when $\theta = 90$ deg at f_0 . A reasonable approximation to this impedance curve over a moderately wide frequency band may be obtained experimentally. Deviations of the experimental impedance curve from the ideal impedance curve result in deviations from the ideal field pattern.

An example of a broad-band crossed-element phased antenna is shown in Fig. 5-28. This antenna consists of a pair of crossed-V-sleeve dipoles mounted approximately a quarter-wave above a ground-plane reflector. The antenna assembly has a coaxial input from which a coaxial line leads to each dipole. At the far end of each of these lines, there is a

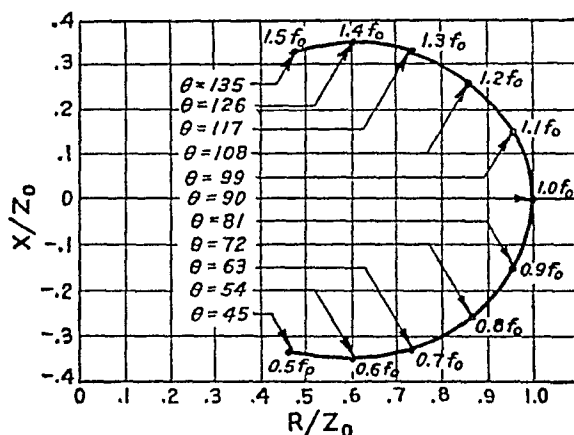


FIG. 5-27.— R - X diagram of ideal input impedance of each element, required to maintain equal power in, and 90-deg phasing between, two elements in space quadrature, over a frequency band where the phasing line is 90 deg long at the frequency f_0 .

type-III balun (see Sec. 3-16). Physically, these two baluns are composed of four parallel tubes arranged symmetrically about the antenna axis. Each balun consists of a diagonally opposite pair of these tubes. The total electrical length of one of the lines is 275 deg at the center frequency of the antenna, and the total electrical length of the other line is 185 deg. Functionally, the phasing line consists of the first 90 deg (at the center frequency) of the longer line measured from the conductor. Its characteristic impedance is 98 ohms. The remainder of the longer line is identical to the shorter and consists of a two-section impedance transformer, as can be seen from the figure. The characteristics of this impedance transformer were chosen so that it would transform the antenna-element impedance curve to a form that more nearly fulfills the requirements of Eq. (5-10) than does this impedance curve itself.

The sleeve structure of the antenna elements is attached to and supported by a square tube that encloses the balun assemblies. The dipoles were made of V form in order to increase the component of radiated field

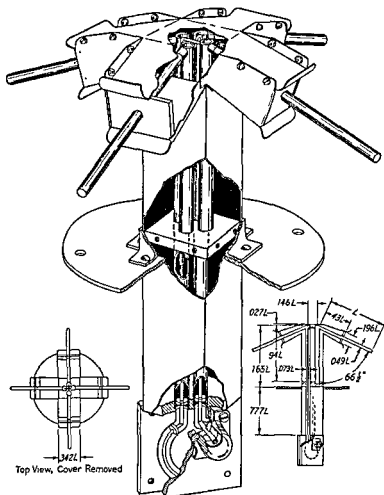


FIG. 5-28 — An example of a broad-band crossed-element phased antenna, with sleeve-dipole elements placed approximately a quarter wavelength above the ground plane, with field pattern single-lobed and approximately circularly polarized in the axis of the lobe.

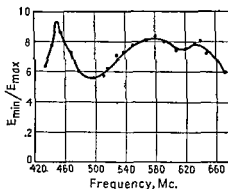


FIG. 5-29.—Ratio of minimum to maximum voltage impressed on a probe dipole in the axis of symmetry of a crossed-element phased antenna as the antenna is rotated.

normal to the ground plane, at directions at a small angle to the ground plane.

The variation of radiated voltage against polarization in a crossed-element phased antenna can be measured by rotating a dipole in the axis of the antenna, or by rotating the antenna, and measuring the voltage impressed on a dipole in the axis of the antenna. Reflection difficulties can be simplified when the test dipole is suspended in the axis of the antenna and the antenna is rotated.

The variation of radiated voltage against polarization for the antenna discussed above stays within reasonable limits between 440 and 660 Mc, a band width of 1.5 to 1. The performance of the antenna may be described graphically by plotting the ratio of the minimum and maximum voltages impressed on a dipole in the axis of the antenna as the antenna is rotated. Figure 5-29 is a graph of these values for the antenna discussed above.

CHAPTER 6

HORNS AND REFLECTORS

By G. STAVIS AND A. DORNE

6-1. Introduction.—The horn and the concave reflector have a property in common that distinguishes them from other types of radiators: both possess definite apertures¹ through which the radiant energy enters free space. The concept of *aperture* may, of course, be profitably applied to any radiating system when the magnitude of the area that it encompasses is an important parameter.

But the aperture is a dominating aspect of the horn and the concave reflector because it exists physically

If the direction of energy flow reverses within a device, it is customarily called a *reflector* (or reflector assembly in order to include the element that illuminates the reflector). If the direction of energy flow through the device is substantially unchanged, it is a *horn*. Usually there is no ambiguity, but this is not necessarily the case. The series of radiators of Fig. 6-1 shows a gradual transition from reflector to horn. They are presented here to emphasize the general similarity between the two types. Some of the examples shown in the figure are not familiar. This does not imply that they might not

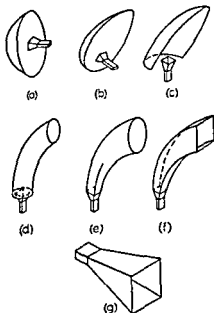


FIG. 6-1.—Some possible arrangements of horns and reflectors.

be useful. In fact, example *c* has proved quite appropriate in some circumstances and is discussed in detail in Sec. 6-8.

Aperture radiators are normally most useful where highly directional radiation patterns are desired. (Slot radiators described in Chap. 7 are an exception.) Since the property of providing highly directional patterns is not particularly manifest unless the aperture is at least half a

¹ The concept of the horn radiator is regarded as sufficiently broad to include the slot antennas of Chap. 7.

square wavelength, horns and reflectors are ordinarily used at high frequencies where structures of this size are not likely to be impracticably large. In practice the aperture radiators employed are frequently much larger than half a square wavelength. Size is not the only consideration that leads to the use of horns and reflectors at microwave frequencies. Their inherent simplicity and freedom from the critical adjustments required by phased multielement arrays make them particularly suitable in microwave applications. These and certain other factors make them particularly adaptable to broad-band work.

In Sec. 1-11 the general relations among aperture, gain, and radiation-pattern shape are discussed. It is clear from that discussion that radiation patterns change gradually with changes in aperture size. Accordingly,

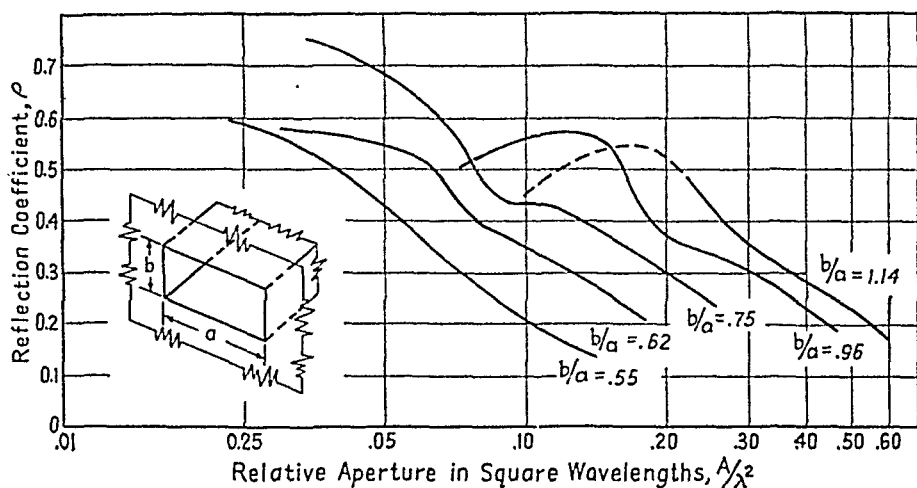


FIG. 6-2.—Reflection at horn aperture vs. aperture dimensions.

an antenna with an aperture of a given physical size has a radiation pattern that changes gradually with frequency, provided that the phase and amplitude distributions of the field across the aperture are relatively constant with frequency. Therefore, in many broad-band applications the design problem consists of illuminating an aperture in a manner that does not vary or varies slowly with frequency. When greater pattern uniformity is required, it can be achieved by varying the illumination so as to compensate for the change in dimensions (in wavelengths) of the aperture. Relatively little work of this type has been done, since the need has not arisen, and the only further reference to the matter will be in Sec. 6-3 where, in connection with the discussion of horns, it will be shown that in some types of horns a compensation of this form automatically takes place.

In general, as is pointed out in Chap. 1, there are at least two regions in an electromagnetic radiator in which discontinuities exist. In the

horn, one discontinuity occurs where the waves enter free space, and another where the guided waves enter the radiator. In a reflector assembly, an additional discontinuity is present where the waves leave the illuminating element. Since the apertures used are likely to be substantially larger than half a square wavelength, the reflection caused by the discontinuity at the aperture is usually negligible. Evidence of this is provided by the curves of Fig. 6-2 which show the ratio of reflected-to-incident-wave amplitudes vs. aperture area, for the case of a $TE_{1,0}$ wave propagated in a rectangular waveguide the open end of which forms the aperture. Several curves are shown, corresponding to guides of various height-to-width ratios. It will be noticed that, although the various curves exhibit many differences, they all show that the reflected wave has a general tendency to decrease with increasing aperture and that for apertures of the order of a half square wavelength the reflection is quite small.

6-2. Methods of Feeding Horns.—In general, horns are fed by means of either waveguides or coaxial transmission lines. In the former case,

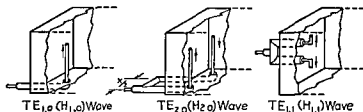


FIG. 6-3—Typical probe arrangements for exciting different types of waves in rectangular guides. (From F. E. Terman, "Radio Engineers' Handbook," McGraw-Hill Book Company, Inc., New York, 1943, Fig. 128, p. 261, by permission)

if the horn is flared at a small angle, and the dimensions of the input end are the same as the dimensions of the cross section of the guide, the reflection that occurs at the input is small. Such a device is inherently broad-band because it effects a gradual transformation of the guided waves into free space waves.

When coaxial transmission lines are used, the transformation is more difficult to effect over a wide band of frequencies. Because it is a convenient procedure that allows some standardization, it has been the practice first to transform from the coaxial line to waveguide. The waveguide is then used to feed any horn. Methods for exciting various modes in rectangular waveguide are shown in Fig. 6-3.¹ These devices all involve the use of probes that couple the coaxial line to the guide, and all are frequency-sensitive since the reflections that accompany the transformation are a function of both the electrical length of the probe

¹ Terman, F. E., "Radio Engineers' Handbook," p. 261, McGraw-Hill Book Company, Inc., New York, 1943.

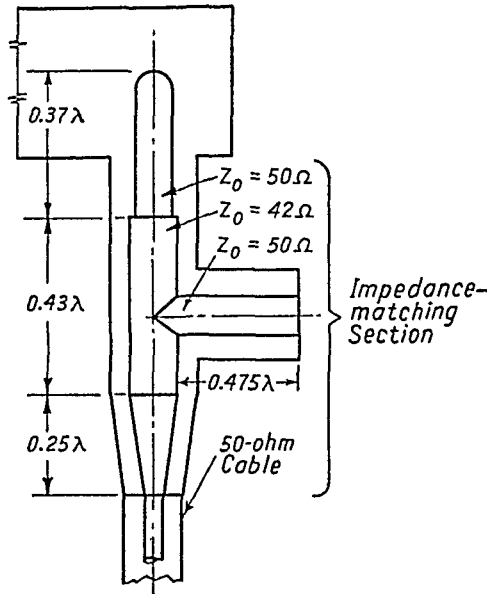


FIG. 6-4.—Coaxial-line-to-waveguide transformer.

and its distance from the back wall of the guide. However, it is possible to compensate for these reflections by the use of suitable matching elements. A transformer so compensated is shown in Fig. 6-4. The voltage standing-wave ratios that are produced on the coaxial line by this transformer are shown in Fig. 6-5. The useful band width of this feed section is about 1.9 to 1.

A type of feed section that does not make use of a probe is shown in Fig. 6-6. This, a relatively reflectionless broad-band transformer of great utility in broad-band practice, is described in detail in Chap. 27 which treats the general subject of coaxial-

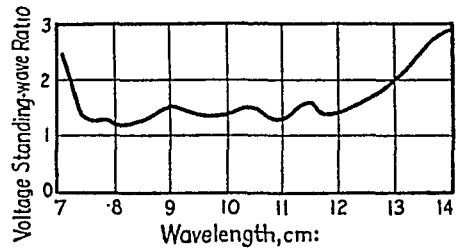


FIG. 6-5.—Voltage standing-wave ratio of waveguide transformer (probe type).

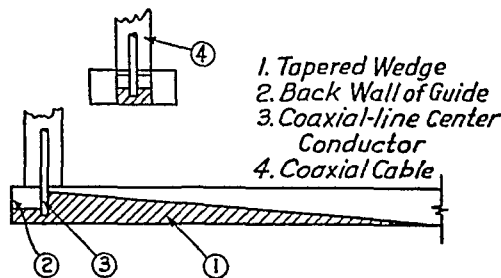


FIG. 6-6.—Tapered type of coaxial-line-to-waveguide transformer.

line-to-guide transformers. For comparison with the curve of Fig. 6-5, a voltage standing-wave-ratio curve for the transformer of Fig. 6-6 is given in Fig. 6-7.

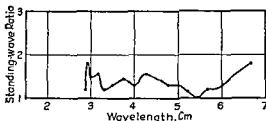


FIG. 6-7.—Voltage standing-wave ratio of waveguide transformer (tapered type).

6-3. Radiation Patterns of Horns.—From the data in Fig. 6-2 it is clear that a horn aperture need not be large to effect a substantially reflectionless transformation from guide to free space. In fact, a flared horn that has been designed to produce a reasonably directional pattern usually requires no modification to meet impedance requirements except when the reflection tolerances are very small. For many purposes, therefore, flared horns may be designed to meet pattern requirements independently of impedance characteristics. One of the simplest and also most useful horns is the straight-sided rectangular horn. This

section is concerned with a discussion of some of its properties.

Let us consider first the sectoral horn shown in Fig. 6-8. This has two parallel sides, two nonparallel sides, and in the usual case a plane aperture. The mode of excitation most commonly used with such horns is the one that corresponds to the $TE_{1,0}$ mode for rectangular guide. In this mode the lines of electric force are perpendicular to, and terminate

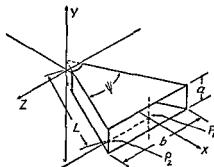


FIG. 6-8.—Sectoral horn

on, the parallel walls of the horn, and the lines of magnetic force are parallel to the parallel walls. The wave front within the horn is cylindrical, and its axis lies at the apex of the horn. Figure 6-9 gives a diagrammatic representation of the field. Chu and Barrow¹ have computed the radiation patterns to be expected from this type of horn. Their results for the patterns through the center of the horn in the plane parallel to the magnetic field (for the region where the pattern is essentially single lobed) are summarized in the curves of Fig. 6-10. This figure shows

¹ CHU, L. J., and BARROW, W. L., Electromagnetic Horn Design, *Trans. A.I.E.E.*, 58, 333 (1939).

beam width vs. aperture for various values of flare angle. It will be noticed that the beam width decreases with increase in aperture and also decreases when the flare angle is decreased. This effect of increasing

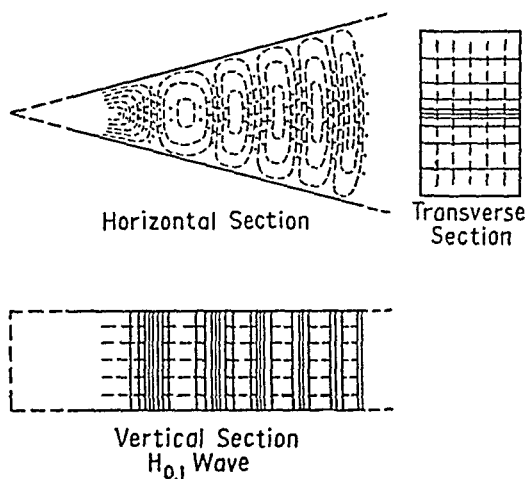


FIG. 6-9.—Field configuration within sectoral horn.

aperture is, of course, not peculiar to the horn, and the general discussion of Sec. 1-11 is completely applicable here. The effect of change of flare angle is a property of the horn itself and results from the cylindrical shape of the wave front mentioned above. The basic point is that, since the wave front in the horn is cylindrical, it would be cylindrical at the aperture if it were not for the distortion at the edges. Thus, the field at the center line of the aperture and the field at the boundary of the aperture differ in phase because of the difference between the path lengths to the two places. Thus for the horn of Fig. 6-8, neglecting the distortion of the field near the edges of the aperture, the field at point P_1 leads the field at point P_2 by $\Delta = 2\pi L/\lambda[1 - \cos(\Psi/2)]$ radians where L is the length of, and Ψ is the angle between, the two nonparallel sides. Since these conditions prevail to a first approximation in actual horns we should expect that the smaller the phase difference Δ , the narrower will be the

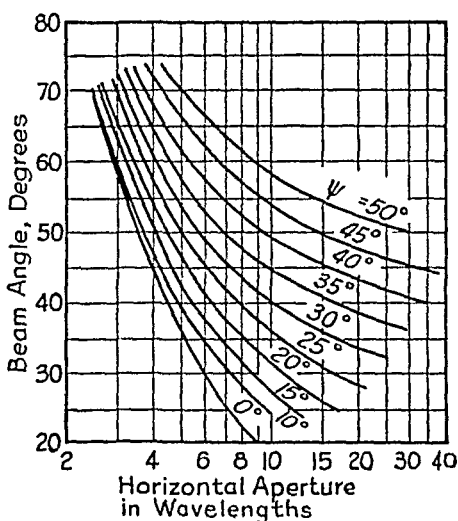


FIG. 6-10.—Beam width between 10 per cent (of field) points vs. aperture for sectoral horn.

radiation pattern. Furthermore, since the phase difference decreases with decrease of Ψ , narrower patterns should be obtained with smaller angle Ψ . Figure 6-10 shows such a trend.

Two conclusions of interest in broad-band practice may be drawn from the foregoing discussion. The first is that because of the relation between radiation pattern and phase difference a horn that has a pattern with a single main lobe at the high-frequency end of its operating band will tend to have patterns with these characteristics at all lower frequencies.¹

The reason for this is that the phase difference between the fields at various points in the aperture will decrease with decreasing frequency, whereas it would have to increase for the main lobe to split. Accordingly,

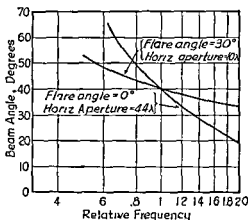


FIG. 6-11.—Curve showing variation in beam width between 10 per cent (of field) points with frequency for horns of zero flare and 30-deg flare.

aperture dimensions and flare angle should be determined on the basis of the highest frequency to be used.

The other conclusion is the previously mentioned fact that constancy of patterns tends to be maintained by virtue of a compensating relation between the changes of aperture size and illumination that occur with frequency. The increase in aperture (in wavelengths) with increase in frequency tends to narrow the pattern, but the increase in phase difference across the larger aperture tends to widen it. The order of magnitude of this compensation can be seen from Fig. 6-11 which shows

¹ The behavior of the lowest order wave in the horn is not the only factor affecting pattern shape. Also important is the relative amplitude of higher order waves which may be present. For many configurations of horns, these higher order modes diminish in amplitude at lower frequencies and therefore produce an effect in the same direction as that described in the above statement. However, when the horn is such that the higher order modes increase in amplitude at lower frequencies, the opposite may be true. For an example of this, see W. L. Barrow and F. D. Lewis, 'The Sectoral Magnetic Horn, *Proc. I.R.E.*, 27, 41 (1939).

two curves (obtained from the data of Fig. 6-10), one of which is for a horn of constant phase (the zero-flare-angle horn) and the other of which is for a horn with a flare angle of 30 deg. The two horns have been chosen to have patterns of the same width (40 deg) at the arbitrary reference frequency. Their apertures are then 10 wavelengths for the horn with the 30-deg flare and 4.4 wavelengths for the horn with 0-deg flare. It is clear that the 30-deg horn shows considerably less variation with frequency.

The preceding discussion has been concerned only with conditions in the plane that lies midway between the two parallel sides of the horn, *i.e.*, the \mathcal{H} plane. With one modification to be mentioned, it is equally applicable to the plane that bisects the horn at right angles to this, *i.e.*, the \mathcal{E} plane. Furthermore, the pattern in one plane is essentially independent of the pattern in the other. The width of the pattern in the \mathcal{E} plane can therefore be determined from the curve for 0-deg flare angle of Fig. 6-10, provided that the width thus obtained is multiplied by a factor of two-thirds to correct for the fact that the field in the \mathcal{E} plane is constant across the aperture, whereas the field in the \mathcal{H} plane varies approximately sinusoidally, in accordance with the usual behavior of $TE_{1,0}$ waves.¹ In general, if the parallel sides of the sectoral horn are diverged so that the horn becomes pyramidal in shape, the foregoing discussion will apply independently to the pattern in either plane.

As previously mentioned, the usual method of feeding a horn is by means of a section of waveguide. If the guide is excited in the $TE_{1,0}$ mode, then the wave in the horn will be of the type we have been discussing. However, since the wave in the guide possesses a plane front, rather than a cylindrical one, it is clear that the wave in the horn immediately beyond the guide cannot have exactly the cylindrical front previously ascribed to it, for if it did the fields would be discontinuous at the junction of horn and guide. Rather the transition from the plane wave of the guide to the cylindrical wave of the horn must be gradual. Therefore, it may be concluded that the curves of Fig. 6-10, calculated upon the basis of a cylindrical wave front, are not applicable to guide-fed horns, unless the horns are sufficiently long for the wave to have become cylindrical at the aperture. If they are not so long it might be expected that the patterns would be sharper than the curves of Fig. 6-10 indicate. The data presented in the following discussion, however, show that an effect of this type is quite manifest even with horns not ordinarily regarded as short. This is of practical significance because many commonly used horns are of these lengths. When very large apertures are required, reflector assemblies are frequently used instead of horns because their bulk is less.

¹ See, for example, Terman, *op. cit.*, Sec. 11.

6-4. Beam Width as a Function of Aperture.—The foregoing statements have indicated the trends in patterns that are to be expected from a consideration of aperture and flare angle. Barrow and Greene¹ have developed expressions for the radiation patterns from a rectangular aperture excited in the $TE_{1,0}$ mode. The assumptions made in obtaining this solution include (1) a plane wave front at the aperture and (2) a sinusoidal distribution of electric field in the \mathcal{H} plane and constant electric field in the \mathcal{E} plane at the aperture. The resultant expressions are as follows:

In the \mathcal{E} plane

$$F_{\mathcal{E}} = C \left[\cos \theta + \sqrt{1 - \left(\frac{m\lambda}{2b} \right)^2} \right] \frac{\sin (\pi W_{\mathcal{E}} \sin \theta)}{\pi W_{\mathcal{E}} \sin \theta} \cos \theta \sin \left(\frac{m\pi}{2} \right) \quad (6-1)$$

In the \mathcal{H} plane

$$F_{\mathcal{H}} = C \left[\cos \theta + \sqrt{1 - \left(\frac{m\lambda}{2b} \right)^2} \right] \frac{\sin \left(\pi W_{\mathcal{H}} \sin \theta + \frac{m\pi}{2} \right)}{(\pi W_{\mathcal{H}} \sin \theta)^2 - \left(\frac{m\pi}{2} \right)^2} \quad (6-2)$$

where F = field intensity

C = constant

θ = angle in the plane under consideration measured from the axis of the horn

m = unity for the mode considered

b = length of the aperture in the \mathcal{E} plane

λ = free-space wavelength

W = ratio of aperture length to λ , in the plane of θ

For large apertures, i.e., for large values of W , it is apparent that the factors

$$\left[\frac{\sin (\pi W_{\mathcal{E}} \sin \theta)}{\pi W_{\mathcal{E}} \sin \theta} \right]$$

in Eq. (6-1) and

$$\left[\frac{\sin \left(\pi W_{\mathcal{H}} \sin \theta + \frac{m\pi}{2} \right)}{\left(\pi W_{\mathcal{H}} \sin \theta + \frac{m\pi}{2} \right) \left(\pi W_{\mathcal{H}} \sin \theta - \frac{m\pi}{2} \right)} \right]$$

in Eq. (6-2) will vary rapidly with θ . The first nulls in the pattern are produced when these factors become zero. In order to obtain the angle between the first nulls in the \mathcal{E} plane, we find the value of θ .

¹ BARROW, W. L., and GREENE, F. M., Rectangular Hollow-pipe Radiators, *Proc. I.R.E.*, 26, 1498 (1938).

$$\frac{\sin (\pi W_{\varepsilon} \sin \theta)}{\pi W_{\varepsilon} \sin \theta} = 0$$

whence

$$\pi W_{\varepsilon} \sin \theta = \pi$$

Since the angle θ_{ε} between first nulls in the ε plane is 2θ ,

$$\theta_{\varepsilon} = 2 \sin^{-1} \frac{1}{W_{\varepsilon}} \quad (6-3)$$

Similarly in the \mathcal{H} plane

$$\frac{\sin \left(\pi W_{\mathcal{H}} \sin \theta + \frac{m\pi}{2} \right)}{\left(\pi W_{\mathcal{H}} \sin \theta + \frac{m\pi}{2} \right) \left(\pi W_{\mathcal{H}} \sin \theta - \frac{m\pi}{2} \right)} = 0$$

whence

$$\pi W_{\mathcal{H}} \sin \theta + \frac{m\pi}{2} = 2\pi$$

and the angle between first nulls in the \mathcal{H} plane is

$$\theta_{\mathcal{H}} = 2 \sin^{-1} \frac{3}{2W_{\mathcal{H}}} \quad (6-4)$$

In Fig. 6-12 a comparison is made between the patterns given by Eqs. (6-1) and (6-2) and experimental results obtained with a particular horn (that of Fig. 6-15). The experimental patterns are similar in shape to the theoretical patterns obtained on the assumption of a plane wave front, but are broader in both the ε and \mathcal{H} planes, indicating that the wave front is somewhat curved; however, the magnitude of the error will be small for short horns, and the equations furnish a practical basis for design of horn radiators.

It is not usual in practice to specify beam width in terms of the angle included between nulls, chiefly because of the difficulty of determining the exact position of the null

when taking measurements. The angle included between the half-power points of the pattern is a more practical measure of beam width, but it is difficult to evaluate the half-power beam width from Eqs. (6-1) and (6-2) without a simplification of the expressions.

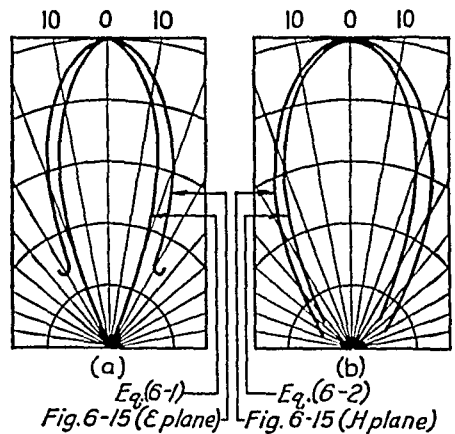


FIG. 6-12.—A comparison of the functions of Eqs. (6-1) and (6-2) with the experimental curves of Fig. 6-15 for $W_{\varepsilon} = W_{\mathcal{H}} = 1.99$.

Let us assume, therefore, that the angle between half-power points β is proportional to the angle between nulls. Also let us make the approximation that the sine of an angle is proportional to the angle. (For

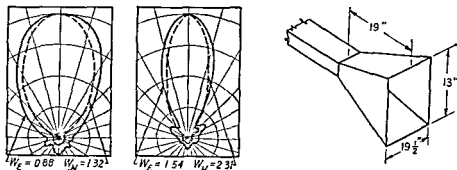


FIG. 6-13.—Horn patterns.

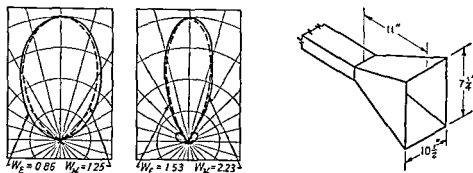


FIG. 6-14 —Horn patterns.

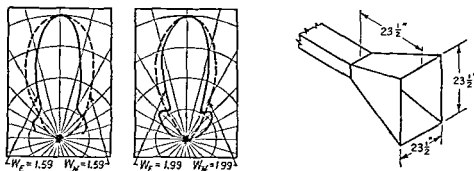


FIG. 6-15.—Horn patterns.

angles less than about 45 deg, the error involved is small.) Then we may write from Eqs. (6-3) and (6-4)

$$\beta_e = \frac{K_e}{W_e} \quad \text{deg} \quad (6-5)$$

$$\beta_{\mathcal{H}} = \frac{K_{\mathcal{H}}}{W_{\mathcal{H}}} \quad \text{deg} \quad (6-6)$$

where $\beta_{\mathcal{E}}$ and $\beta_{\mathcal{H}}$ are the angles between half-power points in the \mathcal{E} and \mathcal{H} planes, respectively, and $K_{\mathcal{E}}$ and $K_{\mathcal{H}}$ are constants.

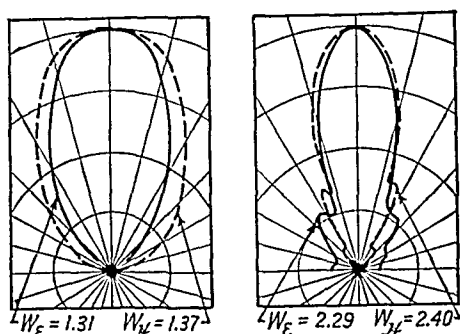


FIG. 6-16.—Horn patterns.

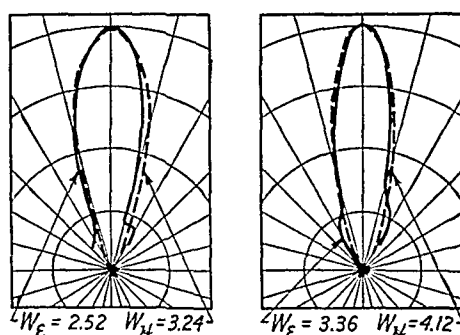


FIG. 6-17.—Horn patterns.

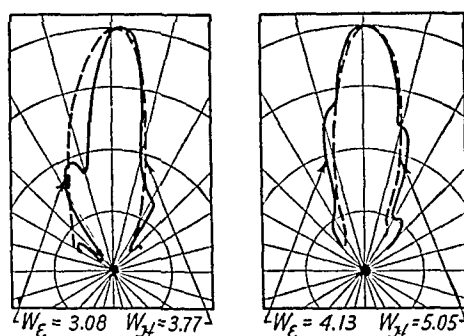


FIG. 6-18.—Horn patterns.

The experimental data for the horns of Figs. 6-13 to 6-19, summarized in Table 6-1, afford a means of determining the values of the constants and also the degree of accuracy to be expected from the use of Eqs. (6-5) and (6-6). Since W and β are known for each of the horns listed,

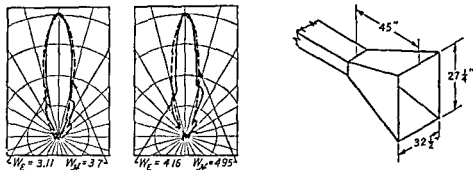


FIG. 6-19.—Horn patterns.

TABLE 6-1.—SUMMARY OF MEASURED DATA FOR THE HORNS OF FIGS. 6-13 TO 6-19

Horn of figure	Dimensions of aperture in λ		Flare angles in E and H planes		Beam widths		Values of K		Frequency range, Mc
	W_e	W_x	ψ_e	ψ_x	β_e	β_x	K_e	K_x	
6-13	0.88	1.32	26	31	56	47	49.5	62	800-1400
	1.1	1.65			49	41	54	68	800-1400
	1.39	2.07			37	31	51	64	800-1400
	1.54	2.31			32	27	49.5	62.5	800-1400
6-14	0.86	1.25	26	29	54	47	46.5	61.2	1400-2500
	1.10	1.6			46	41	50.6	64	1400-2500
	1.35	1.96			37	31	50	68.6	1400-2500
	1.53	2.23			30	27	46	62.5	1400-2500
6-15	1.59	1.59	44	32.5	37	47	58.8	74.7	800-1000
	1.99	1.99			30	37	59.7	73.7	800-1000
6-16	1.31	1.37	45	33	44	52	57.7	71.2	800-1400
	1.63	1.71			36.5	40.5	29.5	69.5	800-1400
	2.07	2.17			26	30	54	65	800-1400
	2.29	2.40			23.5	28.5	54	68.5	800-1400
6-17	2.52	3.24	29.2	32	24.5	23	61.8	74.5	1350-1800
	2.94	3.60			20	19	58.8	68.6	1350-1800
	3.36	4.12			20	18	67.2	74.3	1350-1800
6-18	3.08	3.77	37	42	23	16	1350-1800
	4.13	5.05			23	18.5	1350-1800
6-19	3.11	3.70	30	33	20	16	62.2	62.5	1350-1800
	3.63	4.33			16	15	57.2	65	1350-1800
	4.16	4.95			17	13	70.7	64.3	1350-1800

we may compute an average value of K_ε and $K_{3\varepsilon}$. Using these empirical values of K_ε and $K_{3\varepsilon}$, Eqs. (6-5) and (6-6) become

$$\beta_\varepsilon = \frac{55.8}{W_\varepsilon} \quad (6-7)$$

$$\beta_{3\varepsilon} = \frac{67.2}{W_{3\varepsilon}} \quad (6-8)$$

In order to determine qualitatively what degree of accuracy is afforded by these expressions, calculated beam width is plotted as a function of relative aperture in Figs. 6-20 and 6-21, and the points corresponding to

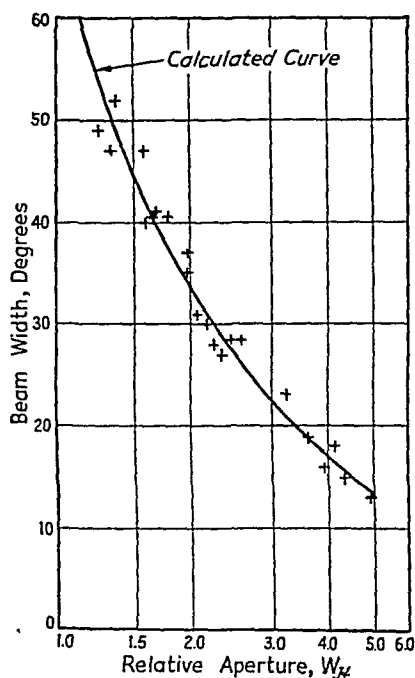


FIG. 6-20.—Beam width vs. relative aperture in 3ε plane.

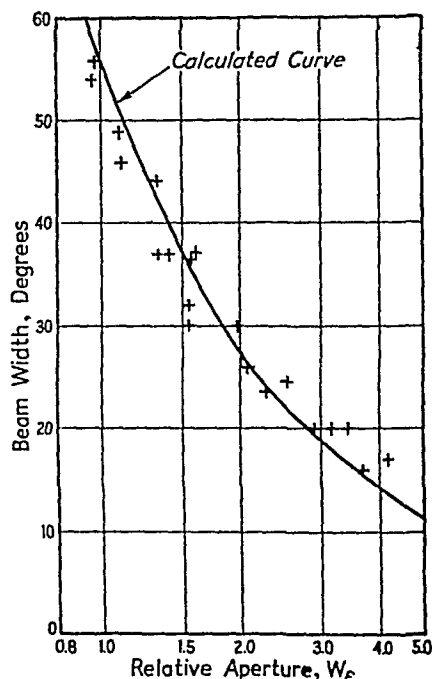


FIG. 6-21.—Beam width vs. relative aperture in ε plane.

the individual beam widths measured for each horn are plotted for comparison. These curves indicate that Eqs. (6-7) and (6-8) give a sufficiently close approximation to the beam width to be of use in the design of a horn radiator.

The horns from which the data of Table 6-1 have been compiled are relatively short in terms of wavelength. Therefore, as pointed out earlier in this section, the field within the horn will not have assumed the distribution pictured in Fig. 6-9, but will be practically plane for small flare angles (less than about 40 deg). Thus, the effect of changing flare angle may be expected to be small, and this is indicated by the curves of Figs. 6-20 and 6-21 which show the points for the various horns. Moreover,

the beam widths obtained are consistent with Eqs. (6-1) and (6-2) which were derived for an aperture illuminated by a plane wave.

If the flare angle is made too large, the wave front at the aperture will be neither plane nor cylindrical but rather some distorted shape, and the pattern of the horn will contain lobes of appreciable magnitude. An example of this behavior is given by the horn of Fig. 6-18. In Fig. 6-19 the patterns for a horn of equal aperture, but less obtuse flare angle, are shown, and here the lobes have been virtually eliminated. This effect of flare angle may be attributed to the transition from waveguide to flared horn where the distortion of the field produces higher order modes of appreciable amplitude. When the flare of the horn is great and the horn is short, these higher order modes are of large amplitude

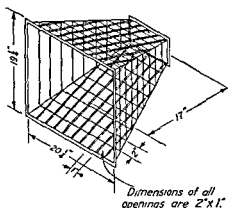


FIG. 6-22.—Lattice-work horn

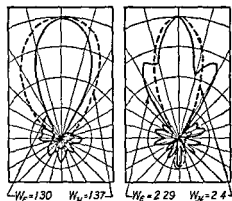


FIG. 6-23.—Patterns showing effect of excessively large mesh openings in lattice-work horn.

and they are not attenuated very rapidly so that they affect the pattern shape by causing an irregular wave front at the aperture. However, as the flare approaches zero length, the case of open-ended waveguide (zero flare angle) is approached and the radiation pattern again will not have significant minor lobes. Thus, it is clear that horn length and flare angle are interdependent in determining radiation-pattern shape.

Sometimes when a horn must be mounted in an exposed location, the windage load upon it is objectionable. In such cases, the windage load may be minimized through the use of open lattice rather than continuous sheet metal for the sides of the horn. A horn constructed in this way is shown in Fig. 6-22. The radiation patterns obtained in the range from 1000 to 1400 Mc are practically identical to those that would be obtained with solid sides on the horn (see Fig. 6-16). However, when the mesh openings become too great, the patterns will be affected. For example, Fig. 6-23 shows the pattern for the identical horn of Fig. 6-22, except that half of the longitudinal brass stringers in the sides parallel to the \mathcal{H} lines

have been removed. The magnitude of the discontinuities is now great enough to produce pattern distortions in the ϵ plane at the higher frequency.

6-5. Circularly Polarized Horns.—There is one type of horn radiator that deserves special attention. It differs from the horns that have been discussed in that it radiates a circularly polarized wave when used for transmitting and will receive signals of any angle of plane polarization as well as circularly polarized signals having the proper direction of rotation. A circularly polarized wave may be produced by an antenna that radiates simultaneously two fields equal in intensity, in which the ϵ vectors are at right angles to each other and are 90 deg out of phase. In general the two fields are not equal and the phase difference between them has an arbitrary value β , rather than 90 deg. If $\epsilon_1 \sin \omega t$ is the electric vector in the horizontal or x -direction and $\epsilon_2 \sin (\omega t - \beta)$ is the electric vector in the vertical or y -direction, the resultant field radiated into space at any instant will be their vector sum. The locus of the end of the resultant vector r can be found as follows:

$$x = \epsilon_1 \sin \omega t \quad (6-9)$$

$$y = \epsilon_2 (\sin \omega t - \beta) = \epsilon_2 (\sin \omega t \cos \beta - \cos \omega t \sin \beta) \quad (6-10)$$

Therefore,

$$\sin \omega t = \frac{x}{\epsilon_1} \quad (6-11)$$

Substitution in Eq. (6-10) gives

$$\frac{y}{\epsilon_2} = \frac{x}{\epsilon_1} \cos \beta - \left(\sqrt{1 - \frac{x^2}{\epsilon_1^2}} \right) \sin \beta \quad (6-12)$$

$$\frac{y^2}{\epsilon_2^2} - \frac{2xy \cos \beta}{\epsilon_1 \epsilon_2} + \frac{x^2}{\epsilon_1^2} = \sin^2 \beta \quad (6-13)$$

Since this is the equation of an ellipse with its center at the origin, a field of this type is said to be *elliptically polarized*. If $\beta = 90$ deg and $\epsilon_1 = \epsilon_2$, then Eq. (6-13) becomes the equation of a circle with center at the origin and radius ϵ_1 . Thus at any time t the resultant r will be equal to ϵ_1 and will have an angular velocity ω . A field of this type is called *circularly polarized*.

In order to produce the requisite space and time quadrature of the field components in a horn radiator the construction shown in Fig. 6-24 may be utilized. The horn is fed from a conventional waveguide through which is transmitted a $TE_{1,0}$ wave, but between the horn and waveguide there is interposed a transition section that displaces the horn about its axis 45 deg from the waveguide. The displacement is such that the wave entering the horn is resolved into two mutually perpendicular compo-

nents, one of which is in the $TE_{0,1}$ mode and the other in the $TE_{1,0}$ mode. Thus, two fields in space quadrature are produced. By causing one component to lag by 90 deg with respect to the other, a circularly polarized field is radiated. This function is performed by the phasing section of the horn, in which the guide wavelength for one mode is made to be different from the guide wavelength for the other so that the phase shift is different. There are two methods of producing this different guide wavelength. One is to make the phasing section rectangular in cross section so that the effective widths and hence the phase velocities are different for the two components. The other is to place a longitudinal sheet of dielectric

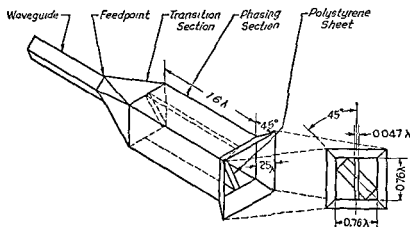


FIG. 6-24.—Circularly polarized horn.

material in the phasing section so that it will affect the velocity of propagation of one of the components only. In this second case the cross section of the phaser may be rectangular, square, or round.

Circularly Polarized Horn without Dielectric.—The phase shift per unit length in the phasing section for each mode is given¹ by

$$\frac{\beta}{l} = \frac{\pi}{a} \sqrt{\frac{4a^2}{\lambda^2} - 1} \quad (6-14)$$

where β = phase shift, radians

l = length of guide, centimeters

a = cross-sectional dimension in centimeters normal to E lines for the particular mode

λ = free-space wavelength, centimeters

¹ This may readily be obtained from the usual expression

$$\frac{\lambda_g}{\lambda} = \frac{1}{\sqrt{1 - \left(\frac{\lambda}{2a}\right)^2}} \quad \text{since} \quad \frac{\beta}{l} = \frac{2\pi}{\lambda_g}$$

Figure 6-25 shows a family of phase-shift curves plotted against wavelength for values of a from 1 to 5. It is possible from curves such as these to design a phasing section that will produce a 90-deg phase shift at a particular frequency. There is an infinite number of possible combinations. Let us assume, for example, that it is desired to operate at a 5-cm wavelength and that the cross section of the phasing section is 3×4 cm. From Fig. 6-25 the phase shift per unit length for the wave having its \mathcal{E} lines normal to the 4-cm side will be about 57 deg per cm. The other mode will undergo a phase shift of about 41 deg per cm. The difference between these is 16 deg per cm. Since the total difference in

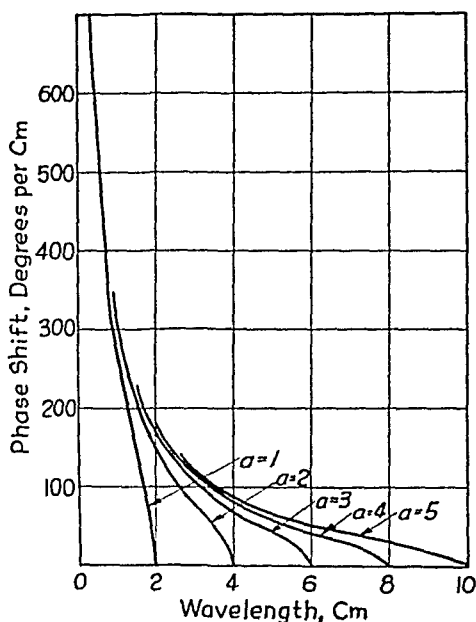


FIG. 6-25.—Phase-shift curves for waveguide with air dielectric.

phase between the two modes is to be 90 deg through the phasing section, the length of the section must be $\frac{90}{16}$, or 5.62, cm.

Examination of the curves shows that if the frequency is changed then the total phase difference between the two modes will no longer be 90 deg, since the phase shift per unit length for these two curves varies markedly with wavelength.

If the phase shift through the phasing section is not 90 deg, then the polarization of the radiated signal will not be circular but elliptical, as indicated by Eq. (6-13). The amount of ellipticity that can be tolerated will determine the band width. If an ellipse such that the ratio of major to minor axis is about equal to 1.4 is acceptable, a band width of about 1.3 to 1 can be obtained.

Circularly Polarized Horn with Dielectric.—The other method of achieving unequal phase shifts per unit length for the two modes is, as mentioned above, to place a sheet of dielectric material in the phasing section so as to influence the velocity of propagation of one mode and not the other. This is illustrated in Fig. 6-24. Here the mode having \mathcal{E} lines parallel to the dielectric sheet will undergo a greater phase shift per unit length than the mode having \mathcal{E} lines normal to the sheet. If the sheet is thin, the velocity of propagation for the wave with \mathcal{E} lines normal to it will be practically unaffected by its presence.

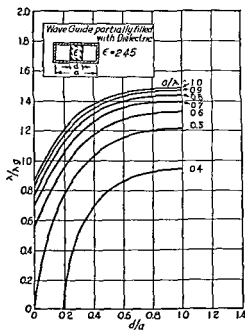


FIG. 6-26—Curves showing effect of polystyrene dielectric on guide wavelength. (From "Microwave Transmission Design Data," Publication No. 23-80, Sperry Gyroscope Company, Inc., Fig. VIII-7, by permission.)

The effect of such a sheet in a waveguide excited by a $TE_{1,0}$ mode with \mathcal{E} lines parallel to it can be seen from Fig. 6-26, which shows the effect on the guide wavelength of sheets of polystyrene of varying thickness. The ratio of free-space wavelength to guide wavelength, which is plotted as ordinate, if multiplied by $2\pi/\lambda$ will give the phase shift per centimeter. Therefore, if a guide is selected with the dimension a equal to 3 cm, we can plot in Fig. 6-27 (as in Fig. 6-25) a family of curves showing phase shift per centimeter as a function of wavelength for different thicknesses of polystyrene sheet.

In this case the difference in ordinates from one curve to an adjacent

curve remains far more constant than was the case with the curves discussed in the preceding section in which the guide dimensions alone produced the phasing. Thus, if the phasing section is made to have a square cross section and the difference in guide wavelength is secured solely through the use of a polystyrene sheet, the band width may be expected to be considerably greater than it would be for the rectangular phasing section without dielectric. In actual experience the band widths that are obtained are more than twice those obtained without dielectric.

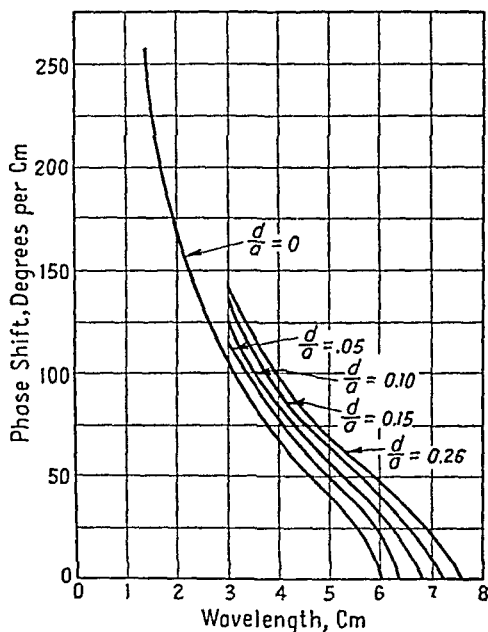


FIG. 6-27.—Phase-shift curves for waveguide with dielectric.

It will be noticed in Fig. 6-24 that the dielectric material is tapered so that the sheet has V-notched ends. This has been done to make the transition from waveguide without dielectric to the phasing section with dielectric a gradual one in order to minimize reflections.

6-6. Measurements of Circularity.—A setup similar to the one shown in Fig. 6-28 has been used to measure the polarization of the radiation from horns of this type. The two horns shown are mounted coaxially. The horn under test is fixed, and the receiving horn, capable of plane-polarization response only, is placed in a rotating mount and rotated in small angular steps about its axis, while the receiver meter is read at each step.

This method of measurement will not give the polarization pattern directly, but will provide data from which the pattern can be deduced.

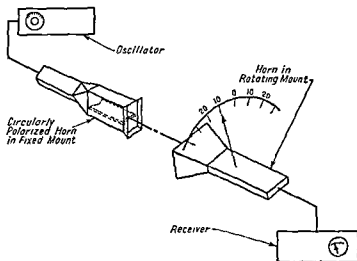


FIG. 6-28.—Setup for measuring polarization patterns

This statement may be clarified by Fig. 6-29, which represents the conditions in a plane normal to the horn axis. The ellipse shown represents a typical polarization pattern of an elliptically polarized wave. The radial dotted lines indicate various attitudes of a receiving horn that is plane polarized.

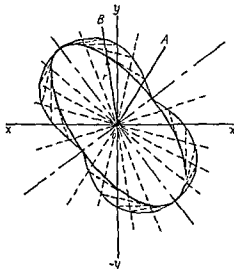


FIG. 6-29.—Diagram illustrating relation of measured dumbbell pattern to actual elliptical polarization pattern.

Since a dipole or plane-polarized horn will respond only to the component of the electric field parallel to its plane of polarization, the maximum voltage received by the receiving horn in any position will occur at the instant at which the projection of the rotating vector \vec{r} (the locus of which is the ellipse) upon the plane of polarization OA of the receiving horn is a maximum. The maximum projection of the rotating vector for each position of the horn will occur when the normals to the dotted lines are tangent to the ellipse. Thus, the envelope of the projections on the plane of polarization of the receiv-

ing horn will give a dumbbell as shown. Figure 6-30 shows three patterns that were measured for a circularly polarized horn. Figure 6-30a, which shows the greatest deviation from circularity, illustrates the dumb-

bell shape of the measured pattern. Since the major and minor axes of the ellipse and dumbbell are equal and coincident, the actual deviation from circularity can be obtained directly from the measured pattern.

More than just the deviation from circularity can be deduced from an inspection of the measured dumbbell, however. In Fig. 6-31 the x -axis

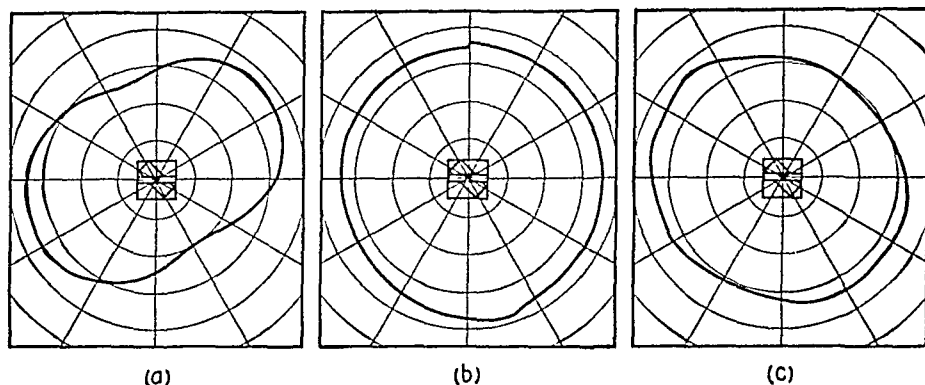


FIG. 6-30.—Circularity patterns.

is taken normal to the polystyrene sheet as a reference and the elliptical pattern is shown in relation to the horn. It is obvious that if the x - and y -intercepts of the dumbbell are equal, then ϵ_1 and ϵ_2 , which are the components of the electric vector in the x - and y -directions, are equal in magnitude and θ , the angle between the major diameter and the x -axis,

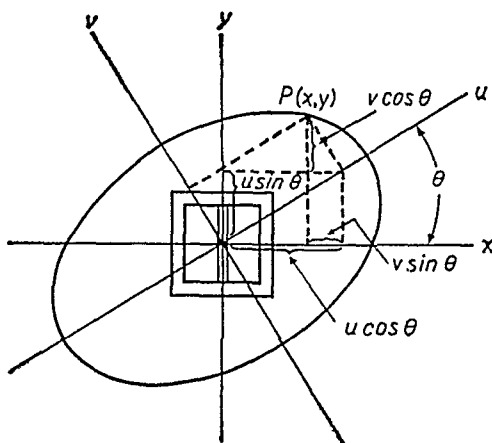


FIG. 6-31.—Rotation of coordinate axes.

will be 45 deg. Furthermore, in this case, θ will be independent of the relative time phase of ϵ_1 and ϵ_2 . If the phase difference β between ϵ_1 and ϵ_2 is 90 deg, but ϵ_1 does not equal ϵ_2 , then the ellipse will lie with its major axis along either the x - or y -axis, depending on whether ϵ_1 or ϵ_2 is greater.

In general, however, it is probable that both conditions (where θ is not equal to 45 deg and the x - and y -intercepts are unequal) will be encountered, in which case it is not obvious what β and the ratio of ϵ_1 to ϵ_2 are. In order to obtain a solution for β and the ratio ϵ_1 to ϵ_2 in terms of θ and the ratio of major to minor axis of the ellipse, let us refer to Fig. 6-31, which shows the elliptical pattern with respect to the orientation of the horn. If we rotate the coordinate axes through the angle θ so that the major axis of the ellipse will be coincident with the new axis of abscissas, the following substitutions must be made in Eq. (6-13):

$$x = u \cos \theta - v \sin \theta \quad (6-15)$$

$$y = u \sin \theta + v \cos \theta \quad (6-16)$$

where the new axes are the u - and v -axes.

Thus, Eq. (6-13) becomes

$$\begin{aligned} u^2 \left[\frac{\epsilon_1}{\epsilon_2} \sin^2 \theta - \sin (2\theta) \cos \beta + \frac{\epsilon_2}{\epsilon_1} \cos^2 \theta \right] \\ + uv \left[\frac{\epsilon_1}{\epsilon_2} \sin (2\theta) - 2 \cos (2\theta) (\cos \beta) - \frac{\epsilon_2}{\epsilon_1} \sin (2\theta) \right] \\ + v^2 \left[\frac{\epsilon_1}{\epsilon_2} \cos^2 \theta + \sin (2\theta) \cos \beta + \frac{\epsilon_2}{\epsilon_1} \sin^2 \theta \right] = \epsilon_1 \epsilon_2 \sin^2 \beta \quad (6-17) \end{aligned}$$

If the coefficient of uv is zero

$$\frac{\epsilon_1}{\epsilon_2} \tan (2\theta) - 2 \cos \beta - \frac{\epsilon_2}{\epsilon_1} \tan (2\theta) = 0 \quad (6-18)$$

from which we get

$$\tan (2\theta) = \frac{2 \frac{\epsilon_1}{\epsilon_2} \cos \beta}{\left(\frac{\epsilon_1}{\epsilon_2} \right)^2 - 1} \quad (6-19)$$

If ellipticity ϵ is defined as the ratio of v -intercept to u -intercept, the square of the ellipticity is seen to be the ratio of the coefficients of the u^2 and v^2 terms and

$$\epsilon^2 = \frac{\left[\frac{\epsilon_1}{\epsilon_2} \sin^2 \theta - \sin (2\theta) \cos \beta + \frac{\epsilon_2}{\epsilon_1} \cos^2 \theta \right]}{\left[\frac{\epsilon_1}{\epsilon_2} \cos^2 \theta + \sin (2\theta) \cos \beta + \frac{\epsilon_2}{\epsilon_1} \sin^2 \theta \right]} \quad (6-20)$$

Equations (6-19) and (6-20) if solved simultaneously for ϵ_1/ϵ_2 give

$$\left(\frac{\epsilon_1}{\epsilon_2} \right)^2 = \frac{\epsilon^2 \tan^2 \theta + 1}{\epsilon^2 + \tan^2 \theta} \quad (6-21)$$

To find β this result may be substituted in Eq. (6-19). Thus, the factors that control the ellipticity may be derived from ϵ and θ which are taken from the measured patterns.

6-7. Reflectors; General Considerations.—Reflectors have a prominent part in the field of broad-band microwave directional antennas. Very narrow patterns can be obtained through their use, with structures in which the front-to-back dimensions are much less than would be those of the equivalent horns. This economy of size is frequently very important and often dictates the use of a reflector. Also the reflector sometimes fulfills special functions not obtainable with a horn. This is true, for example, of the antenna described in Sec. 6-8, which permits pointing of the antenna beam in any horizontal direction, by rotating the reflector alone. The horn that illuminates it remains stationary, and the design is simplified by the removal of the necessity for a broad-band rotating waveguide joint.

An aperture will produce a smooth essentially single-lobed radiation pattern, if its illumination meets the following two conditions:

1. The intensity of the field in the aperture increases gradually from zero at the edges to a maximum value at the center.
2. The phase of the field in the aperture is constant.

Actually, as is evident from the previous discussion of horns, these conditions may be departed from by substantial amounts, without either causing the patterns to become irregular or greatly increasing the amplitude of the minor lobes. The amount of pattern distortion that will accompany a particular deviation from one of the conditions will depend upon how closely the other condition is met, and also upon the dimensions of the aperture. Sometimes, of course, the conditions are deliberately departed from in order to produce a type of pattern distortion, *e.g.*, widening of the beam, that under the particular circumstances is desirable. Nevertheless, the most usual objective is to meet the two conditions as nearly as possible.

One of the most commonly used reflectors is the paraboloid of revolution. As is well known, such a surface converts the energy that impinges upon it from a source of spherical waves located at its focus into a plane wave in the aperture of the reflector. In general practice it is fairly common to use a *primary radiator*, *i.e.*, the radiator at the focus, that has a directional pattern that points into the parabola. The purpose is to avoid as nearly as possible the transmission of energy directly from the primary radiator into space since this radiation would not be focused by the reflector and hence would not obtain the desired directional characteristics. In broad-band work this is a particularly desirable practice because the direct wave, if it exists, will be at a phase with respect to the

reflected wave that changes with frequency. At some frequencies it will add to the reflected wave, and at others it will subtract. The rate at which this occurs will depend upon the focal length of the parabola in wavelengths, but in any case the uniformity of patterns with frequency will be decreased.

It may be deduced from the above that the primary radiator for a paraboloidal reflector should produce a directional pattern the wave front of which is spherical. If maximum radiation is directed toward the center of the reflector, and the radiation diminishes to a negligible value at the edges, the two conditions originally set forth will in general be met. Suitable primary radiators may consist either of dipole assemblies or horns. A broad-band example of the former type which has been successfully used is the sleeve-and-reflector assembly of the general type described in Sec. 5-6. A modification of this is described in connection with direction-finder antennas in Sec. 10-31. It should be noted that an antenna of this type does not actually produce a spherical wave. However, if the solid angle subtended by the reflector is sufficiently small, the waves in the region of the reflector will be approximately spherical in shape, and satisfactory patterns can be obtained. In general this condition can be met by making the focal length of the paraboloid large compared with either the dimensions of the primary radiator or the dimensions of the aperture. Under this circumstance the primary radiator will be located outside the aperture, and its pattern must be highly directional in order that all its radiation shall impinge upon the reflector. It should be noticed that the directional pattern involved here is the pattern of the primary radiator which exists at distances about equal to the separation between the primary radiator and the reflector. In general, this pattern is not the same as the pattern at a great distance from the primary radiator. For example, consider two horns, either of which might be used as a primary source. Suppose that one of them has a much larger aperture than the other and its pattern (in the usual sense) is by far the narrower of the two. At sufficiently small distances from the horn the pattern due to the smaller aperture will be the sharper. This is immediately evident if we consider the extreme case of the distance equal to half the width of the larger aperture. Then with the larger horn the entire path along which measurements are made will lie within the beam of the horn, whereas with the smaller horn only the center of the path will be in the beam and the measured pattern will necessarily be narrower. Accordingly primary radiator patterns must be determined at the distance that is equal to the separation that will exist between the primary radiator and the reflector to be used. It is to be expected that the primary pattern of a radiator will, in general, vary considerably with frequency.

In connection with primary patterns, it may also be noted that the effective aperture of the paraboloid is primarily determined by the primary radiator. For example, if the radiated pattern is to consist of a beam of circular cross section, then the primary pattern should possess radial symmetry. On the other hand, if the lobe of the radiated pattern is to be broader in one direction than in the other, it is necessary that the primary pattern be such that the illuminated portion of the aperture is also broader in one direction than in the other. However, since the smaller effective-aperture dimension will correspond to the broader radiated pattern, the primary pattern should be broad in the plane in which the radiated pattern is to be narrow, and vice versa.

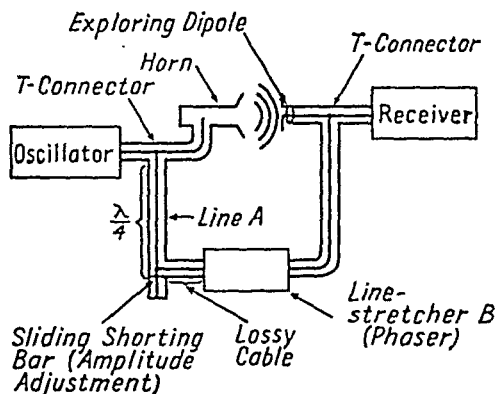


Fig. 6-32.—Schematic drawing of measurement setup for determining equiphase surface of horn radiator.

It is possible, at least in the case of horn primary radiators, to design them so that the front of the wave that they radiate is quite spherical, even at very small distances. An example of a horn in which this has been accomplished through adjustment of the length and angle of its sides is shown in Fig. 6-24. This horn has been used in conjunction with the reflector described in the next section. It was designed empirically through the use of the wave-front measuring setup shown in Fig. 6-32. Essentially the measurement procedure consists of comparing the phase at various points in the field with respect to the phase of a reference signal injected through transmission line from the signal generator to the receiver. Moving the short circuit on line A will adjust the amplitude of the voltage transmitted to the receiver but will have negligible effect on the phase of this energy, provided that the length of line A is not permitted to vary much from the quarter wavelength. The phase of the signal directly transmitted by cable to the receiver is controlled by the line stretcher B. Because of the lossy cable the receiving dipole is effectively isolated from the transmitter. The procedure, briefly, consists of setting the dipole at a fixed point in front of the horn radiator

and adjusting the line stretcher and short-circuiting bar for zero signal, as indicated by the receiver. After this initial adjustment the line stretcher is not disturbed. The exploring dipole is then moved about so as to search for points at which nulls are detected. It may be necessary occasionally to adjust the short-circuiting bar to equalize the amplitudes of signals arriving via the two paths. The locus of the dipole positions at which these nulls are found will be an equiphase surface of the wave.

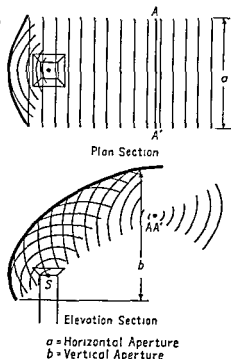


FIG. 6-33.—Plan and elevation sections showing generation of cylindrical equiphase front by horn and reflector of appropriate shapes. (a) horizontal aperture; (b) vertical aperture.

tion, we should change the phase distribution in that direction, but leave it unchanged in the direction at right angles. Thus, we might allow the phase along a horizontal line through the center of the reflector aperture to be unchanged, but delay the phase at other points of the surface by amounts proportional to their distance from this center line. Another possibility is to produce a phase distribution such that the new equiphase surface would be a section of a horizontal circular cylinder (see Fig. 6-33). In this case the vertical directivity would be a function of the radius of the horizontal cylinder generated at the aperture as well as of the vertical aperture dimension, and the horizontal directivity would be a function of the horizontal aperture dimension. The wave at

In general, except when special considerations preclude its use, the horn appears to be a more suitable primary radiator than the dipole assembly. Its radiation characteristics are more readily controlled, and its construction offers fewer problems at high frequencies.

6-8. Elpar (Elliptical-paraboloid) Reflector.—It is often desirable to secure a space pattern that is not circular in cross section, but that has different beam widths in the horizontal and vertical planes. As mentioned above, this may be achieved with a properly illuminated paraboloid. In this case the wave front at the aperture is plane, and the beam shape is controlled by the aperture dimensions. It may also be achieved by appropriately departing from the condition of a plane wave front at the aperture of the reflector. Specifically, to broaden the pattern in one direc-

the aperture would be similar to that which exists with a sectoral horn whose parallel sides are vertical.

The method of determining the theoretical shape of the reflector required to produce such a cylindrical wave front is given below. Figure 6-33 shows the desired end result where the shape of the reflector is such as to focus the energy from the horn source upon the line AA' . In effect the source of the final pattern will appear to be the line AA' instead of the actual point source at S .

For the purpose of analysis let us construct Fig. 6-34 where P_1 represents the primary source (S in Fig. 6-33) and where the z -axis represents the line focus (AA' in Fig. 6-33). The surface $ABCD$ represents a portion of the reflector and is the locus of the point P_2 such that the distance $(P_1P_2 + P_2Q)$ is constant. This requirement is necessary in order that all the energy originating at P_1 arrive in phase at the z -axis. Expressing this condition in polar coordinates, we have the relation

$$\sqrt{(\rho \sin \theta)^2 + (\rho \cos \theta - \rho_1)^2} + z^2 + \rho = K \quad (6-22)$$

in which the various quantities are defined by Fig. 6-34. This can be rewritten as

$$z^2 = \rho(2\rho_1 \cos \theta - 2K) + (K^2 - \rho_1^2) \quad (6-23)$$

If in this expression θ is a constant, we have the expression in polar coordinates for a parabola that lies in a plane containing the z -axis. In particular if θ equals zero, so that the plane under consideration becomes the one that contains the point P_1 (the location of the point source), Eq. (6-23) becomes

$$z^2 = 2 \left(\rho - \frac{K + \rho_1}{2} \right) (\rho_1 - K) \quad (6-24)$$

This is a parabola whose vertex lies on the x -axis at the distance $(K + \rho_1)/2$ from the origin and whose focal length equals $(\rho_1 - K)/2$. We shall refer to the curve of Eq. (6-24) as the *principal parabola*.

Now let us show that if $z = Z_1$ (a constant) we obtain ellipses in all planes normal to the z -axis.

If in Eq. (6-22) we set $z = Z_1$ we obtain

$$Z_1^2 = \rho(2\rho_1 \cos \theta - 2K) + K^2 - \rho_1^2 \quad (6-25)$$

Now if we change to rectangular coordinates where $x = \rho \cos \theta$, $y = \rho \sin \theta$, and $\rho = \sqrt{x^2 + y^2}$, Eq. (6-25) becomes

$$Z_1^2 = 2\rho_1 x - 2K \sqrt{x^2 + y^2} + K^2 - \rho_1^2 \quad (6-26)$$

After considerable manipulation this becomes

$$\left[x + \frac{\rho_1(-K^2 + \rho_1^2 + Z_1^2)}{2(K^2 - \rho_1^2)} \right]^2 + \frac{K^2 y^2}{K^2 - \rho_1^2} = \left[\frac{K}{2} \left(\frac{-K^2 + \rho_1^2 + Z_1^2}{K^2 - \rho_1^2} \right) \right]^2 \quad (6-27)$$

If we let

$$\frac{-K^2 + \rho_1^2 + Z_1^2}{K^2 - \rho_1^2} = r$$

then we have from Eq. (6-27)

$$\frac{\left(x + \frac{\rho_1 r}{2} \right)^2}{\frac{K^2 r^2}{4}} + \frac{y^2}{\frac{(K^2 - \rho_1^2) r^2}{4}} = 1 \quad (6-28)$$

which is the equation of an ellipse with its center at $x = -(\rho_1 r)/2$, $y = 0$. The general equation for an ellipse is

$$(x - h)^2/a^2 + (y - k)^2/b^2 = 1,$$

where the center is at (h, k) and the distance from the center to either focus is $\sqrt{a^2 - b^2}$. The distance from the center to either focus for the ellipse of Eq. (6-28) is therefore

$$\sqrt{\frac{K^2 r^2}{4} - \frac{K^2 r^2}{4} + \frac{\rho_1^2 r^2}{4}} = \frac{\rho_1 r}{2}$$

But since the center is at $x = -\rho_1 r/2$, then one focus of all the ellipses for all odd values of z must lie on the z -axis and also the other focus for the major ellipse, i.e., for $z = 0$, when $r = -1$, must lie at

$$x = -2\rho_1 r/2 = \rho_1.$$

To sum up, we have shown that the locus of a point P_2 such that $P_1 P_2 + P_2 Q$ is constant is a surface whose sections in planes that contain the z -axis are parabolas and whose sections in planes normal to the z -axis are ellipses. Further, the ellipses will be such

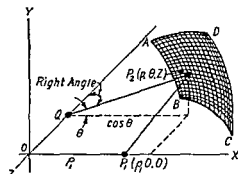


FIG. 6-31.—Space relations for deriving elpar surface.

that one focus will lie on the z -axis, and the other focus of the major ellipse will lie at the source. It can be seen from Eq. (6-28) that the various ellipses have the same ratio of major to minor axis.

The fact that the shape of the surface is such that $P_1P_2 + P_2Q$ is a constant does not assure that AA' be a focal line. It is also necessary to determine whether all rays from the source will be reflected from the surface so as to follow the desired paths. In other words, it must be established that the angle of incidence which the ray P_1P_2 makes with the surface will be equal to the angle of reflection with which P_2Q leaves the surface and that the normal to the reflector surface at every point will lie in the plane of the incident and reflected rays. For convenience, let us redraw Fig. 6-34 as in Fig. 6-35. This time the primary source P_1 is at the origin 0, and the line focus is in the xy -plane at $y = d$ so that their positions are interchanged. The point P_2 as before is on the reflector surface. Again we will set the path length from primary source to line

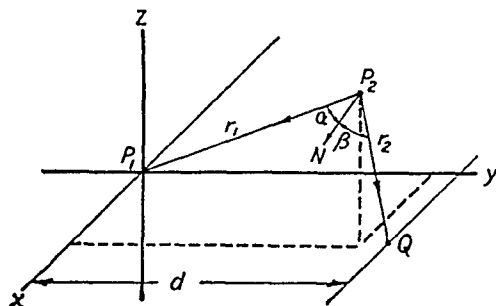


FIG. 6-35.—Space relations for deriving elpar surface.

focus equal to a constant K . In this proof it will be more convenient to work with rectangular coordinates, in which the equation of the surface is

$$K = \sqrt{x^2 + y^2 + z^2} + \sqrt{(y - d)^2 + z^2} \quad (6-29)$$

This is the same surface as that represented by Eq. (6-22) but expressed in rectangular coordinates. First let us obtain a vector normal to the surface at P_2 . For this purpose it is useful to observe that if K is assigned two values K_1 and K_2 that differ by a very small amount one gets two surfaces parallel to each other, just as one gets two equipotential parallel surfaces by assigning two adjacent values of potential. In both cases the direction of the gradient is normal to the surfaces. Therefore the gradient of K gives the direction cosines of the normal.

$$\begin{aligned} \nabla K &= \bar{i}x(x^2 + y^2 + z^2)^{-1/2} + \bar{j}y(x^2 + y^2 + z^2)^{-1/2} \\ &\quad + \bar{j}(y - d)[(y - d)^2 + z^2]^{-1/2} \\ &\quad + \bar{k}z(x^2 + y^2 + z^2)^{-1/2} + \bar{k}z[(y - d)^2 + z^2]^{-1/2} \end{aligned} \quad (6-30)$$

Let

$$L_1 = (x^2 + y^2 + z^2)^{-1/2}$$

and

$$L_2 = [(y - d)^2 + z^2]^{-1/2}$$

Then Eq. (6-30) becomes

$$\nabla K = iL_1x + j[L_1y + L_2(y - d)] + k(L_1z + L_2z) \quad (6-31)$$

Now the unit vector in the direction of P_1P_2 is

$$\bar{r}_1 = -(ix + jy + kz)L_1$$

and the unit vector in the direction of P_2Q is

$$\bar{r}_2 = [j(d - y) - kz]L_2$$

Now if $\nabla K \times \bar{r}_1 = \bar{r}_2 \times \nabla K$, then the normal to the reflector surface lies in the plane containing \bar{r}_1 and \bar{r}_2 . This can readily be shown to be true since

$$\nabla K \times \bar{r}_1 = -L_1 \begin{vmatrix} i & j & k \\ L_1x & L_1y + L_2(y - d) & L_1z + L_2z \\ x & y & z \end{vmatrix} \quad (6-32)$$

$$\nabla K \times \bar{r}_1 = -L_1L_2[-izd + jzx - kx(y - d)] \quad (6-33)$$

and similarly

$$\bar{r}_2 \times \nabla K = -L_1L_2[-izd + jzx - kx(y - d)] \quad (6-34)$$

and the right-hand member of Eq. (6-32) is the same as that of Eq. (6-34).

Furthermore, if $\nabla K \times \bar{r}_1 = \bar{r}_2 \times \nabla K$, then the angle included between ∇K and \bar{r}_1 must be equal to the angle included between ∇K and \bar{r}_2 , so that the angle of incidence must be equal to the angle of reflection.

Thus we have derived the shape of the surface required for obtaining a line focus and have proved that the energy from the point source will indeed be reflected in the desired directions provided that the laws of optical reflection are applicable.

The beam width in the horizontal direction will be a function of the length of the line focus that is illuminated. The vertical directivity will be a function of the primary-source beam width and the orientation of the reflector with respect to the primary source. This is illustrated in Fig. 6-36 where a horn of beam width α is used as a primary source. The major ellipse of the reflector is pictured in various attitudes with respect to the horn. The resultant beam width β (after reflection) changes with the orientation of the reflector with respect to the horn.

An antenna that was designed in accordance with the above principles is shown in Fig. 6-37. Here the axis of the horn is in the plane of the major ellipse, but at an angle of 60 deg to the axis of the major parabola.

A unique feature of this antenna is the means of rotating the direction of the beam in the horizontal plane. Here the reflector can be rotated

about the axis of the horn while the horn remains stationary, provided that

1. The axis of the horn is vertical
2. The horn pattern is symmetrical about the horn axis
3. The illumination of the reflector both in intensity and polarization is constant as the reflector rotates about the horn axis

This last condition is satisfied when a circularly polarized horn is used for a primary source.

It is evident that the beam produced by an elpar surface can be widely varied by suitable adjustments of the parameters. An additional method, which widens the horizontal beam width of the antenna without producing much effect on the vertical beam width, involves the addition of plane baffles at each side of the elpar surface. Figure 6-37 shows a reflector that incorporates planes of this kind. Experimental evidence indicates that their primary effect is to limit the horizontal length of the aperture, thus widening the horizontal patterns. Undoubtedly, there are also secondary effects consisting of distortions of the wave shape. However, the evidence indicated that effects of this type are small. With the baffles shown in Fig. 6-37, the horizontal patterns were about 1.5 times as wide as they were for the reflector without the baffles. It should be mentioned that when the reflector was used without the baffles, it was necessary to extend it in the horizontal direction in order to collect all the energy from the horn. The horizontal and vertical patterns for the reflector and horn of Fig. 6-37 are given in Fig. 6-38. A point of practical interest in connection with this reflector is the fact that metal-plated molded Fiberglas has proved to be a satisfactory material for its construction.

It may be noted that a similar although simpler type of antenna is described in Sec. 10-40. In this case the reflector consists of a section of a cylindrical paraboloid rather than an elliptical paraboloid. The vertical patterns that are shown in Fig. 10-45 possess secondary lobes of

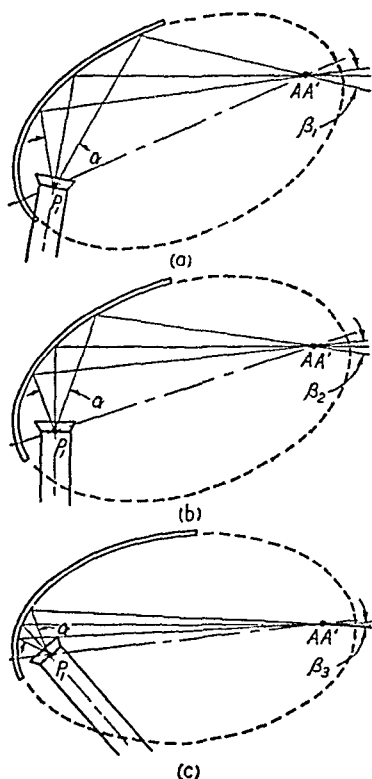
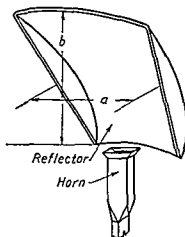


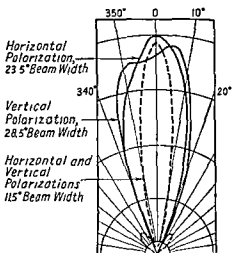
FIG. 6-36.—Diagrams showing dependence of vertical beam width on attitude of reflector.

appreciable magnitude as a result of the erratic phase front produced by this reflector. The shape of the vertical pattern varies considerably with frequency, indicating that phase difficulties are indeed the chief influenc-



a = Horizontal Aperture
 b = Vertical Aperture

FIG. 6-37.—Elpar reflector and horn: (a) horizontal aperture; (b) vertical aperture.



— Vertical Plane
 --- Horizontal Plane

FIG. 6-38—Measured patterns for elpar antenna at $\lambda = 10$ cm.

ing factor in the shape of these patterns. The comparison of the patterns resulting from these two types of reflector surface indicates that the use of the more complex double-curved elpar surface is justified where accurate pattern shape in both vertical and horizontal planes is important.

CHAPTER 7

SLOT ANTENNAS

BY A. DORNE AND D. LAZARUS

7-1. Introduction.—As was discussed in Sec. 1-3, a slot antenna in a flat ground screen may be regarded as a section of waveguide that is terminated in a horn of 180-deg flare angle. It is frequently useful also to consider the more general concept of the slot antenna as an *aperture* that transforms guided waves to free-space waves.

Electromagnetic waves can be introduced at the aperture of the slot by a variety of methods. Generally the frequency sensitivity and behavior of the antenna will depend to a great extent on the method of

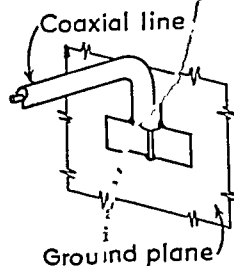


FIG. 7-1.—Coaxial-line-fed slot.

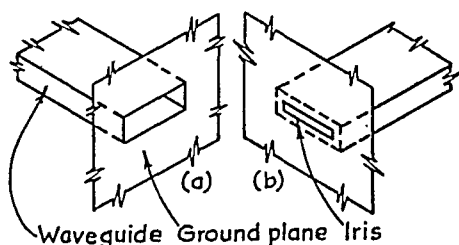


FIG. 7-2.—(a) Waveguide-fed slot; (b) waveguide-fed slot with iris.

excitation of the aperture. One of the simplest methods of setting up waves at a slot is by connecting the individual conductors of a coaxial cable or other two-conductor transmission line to opposite sides of the ground screen immediately around the aperture (see Fig. 7-1). A second obvious method of feeding the slot is by means of a section of waveguide, as in Fig. 7-2. In this case, the slot may consist merely of the direct termination of the waveguide in a ground plane, as in Fig. 7-2a, or may be an iris placed across the aperture of the guide, as in Fig. 7-2b. A third method, shown in Fig. 7-3, is that in which the slot is cut in the surface of the transmission line or waveguide. With this system, waves can be introduced at the aperture either by leakage from the field within the line or guide, or by means of a probe extending into the interior of the transmission line. A fourth obvious method of introducing waves at the slot is by incident radiation from other sources. A narrow rectangular slot, used in this manner, serves as a wave filter, since the aperture will

transmit only waves that are polarized at right angles to the long dimension of the slot.

If excited in proper phase, an array of slots may be used to transmit energy in a directive pattern in a manner analogous to that employed with dipoles. Because of the simplicity of construction possible with slot arrays, they are very often used at extremely high frequencies where construction of dipole elements is impractical.

By its nature, a slot antenna introduces a rather abrupt discontinuity in the path of a wave traveling into free space. In general, therefore, some energy will be radiated and some returned into the initial transmission system in the form of a reflected wave. As in conventional transmission-line practice, we may consider the values of current and voltage excited in the transmission system by the radiated and reflected waves

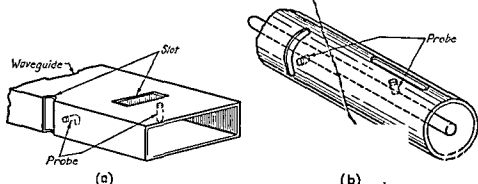


FIG. 7-3.—(a) Probe-fed slots in waveguide; (b) probe-fed slots in coaxial line

and express the performance of a slot antenna in terms of the impedance mismatch introduced on the transmission line feeding the antenna.

Slot antennas may be made broad-band by application of the general principles outlined in Chap. 1. This problem is discussed in some detail in a later section.

One of the most important practical advantages of the slot antenna arises from the fact that the feed section required for excitation of the slot may be placed below the surface of the ground screen in which the slot is cut. Then no protrusion exists on the surface, and since the aperture of the slot itself may be covered with a dielectric sheet, there is no need for even a hole. In the case of aircraft installations, such an antenna has the advantage that it offers no air resistance. In the case of very high-speed aircraft this consideration is particularly important, since the use of conventional antennas introduces serious drag and vibration problems.

7-2. Radiation Patterns of Slot Antennas.—In order to ascertain the type of radiation patterns to be expected with slot antennas, we may first

consider that obtained with the simplest type of slot antenna consisting of an infinite plane metal sheet in which a narrow rectangular slot has been cut. As has been shown in Sec. 1-7, the radiation pattern produced by such a slot is closely related to that which would be produced by a thin flat dipole that would just fill the aperture. The difference between the radiation patterns of the slot and dipole is one of polarization only, since it has been demonstrated that the electric-field distribution obtained with a slot radiator is identical with the magnetic-field distribution in the radiation field of the equivalent dipole. Thus, for a slot that is

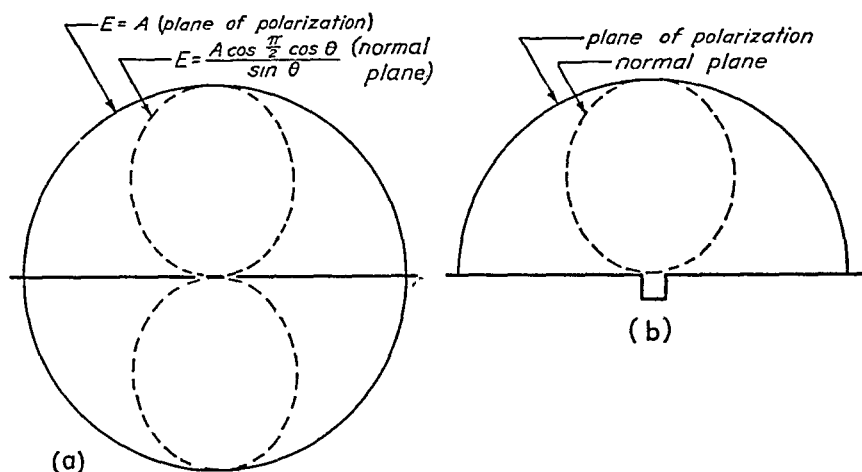


FIG. 7-4.—(a) Radiation pattern of resonant half-wave slot in infinite ground plane; (b) radiation pattern of resonant half-wave slot with cavity behind slot.

exactly $\frac{1}{2}$ wavelength wide,¹ the electric field distribution will be uniform over 360 deg in the plane of polarization and of the form

$$\mathcal{E} = A \left[\frac{\cos(\pi/2 \cos \theta)}{\sin \theta} \right]$$

in the normal plane² (see Fig. 7-4a).

This presumes that the slot is free to radiate on both sides of the ground sheet. If, as in normal use, the radiation to one side is restricted by use of an enclosing cavity behind the slot, the resultant radiation pattern will be just one-half the previous pattern, as shown in Fig. 7-4b. The electric vector in the plane of polarization has a discontinuity at the surface of the ground screen since the tangential component cannot exist at the surface and the normal component is excited only on one side.

¹ Because of the possibility of considering a slot as the termination of a waveguide in a ground plane, it is convenient to refer to the longer and shorter dimensions of the aperture as width and height, rather than length and width.

² See, for example, STRATTON, J. A., "Electromagnetic Theory," Chap. VIII, McGraw-Hill Book Company, Inc., New York, 1941.

In cases where the ground plane is very large, but less than infinite, or is of finite conductivity, the pattern in the plane of polarization will be less than 180 deg wide.

7-3. Effect of Size of Ground Plane on Radiation Patterns.—If the ground plane is less than infinite in the plane normal to the plane of polarization (the \mathcal{H} plane), the size of the ground plane should not have a disturbing effect on the radiation patterns of the slot, since the electric-field vector is tangential to the ground surface in this plane and approaches zero at angles close to the ground plane, even for an infinite plane. If the ground plane is less than infinite in the plane of polarization (the \mathcal{E} plane), an abrupt discontinuity in electric-field distribution will result at the edges of the ground screen, since the normal component of the field

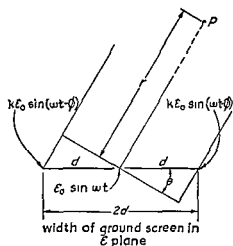


FIG. 7-5.—Ground plane with slot and image radiators.

will be discontinuous at the surface of the ground screen and continuous in the space immediately outside the ground screen. Because of this effect, there will be reflected waves from the edges of the ground screen, which will tend to interfere with the direct wave from the slot. In the solution of this problem, it is convenient to account for the reflected waves at the edges of the ground plane by image sources located at the edges of the ground screen. In Fig. 7-5 the radiating slot is shown at the center of a ground screen of width $2d$. The field is to be computed at a point P far removed from the slot at an angle $\pi/2 - \theta$ with respect to the ground plane. Since $r \gg d$, we may consider the radius vectors drawn from the point P to the center and edges of the ground plane to be parallel.

Let us now postulate the existence of a field of intensity $E_0 \sin \omega t$ at the center of the ground sheet and two equal image fields of intensity $kE_0 \sin (\omega t - \phi)$ at the edges. The angle ϕ is the phase angle of the image radiators and is a function of the distance d from the center of the ground screen to the edges. ϕ will remain fixed for any given value of d . The factor k is an amplitude constant much less than unity, which is also related to d , since the intensity of the image fields must decrease as d increases.

Then the field at point P is given by

$$E_p = E_{p0} [\sin \omega t + k \sin (\omega t - \phi - \beta d \sin \theta) + k \sin (\omega t - \phi + \beta d \sin \theta)]$$

where ε_{p0} is a constant depending on ε_0 and r and the phase-shift constant $\beta = 2\pi/\lambda$. Expanding trigonometrically, we obtain

$$\frac{\varepsilon_p}{\varepsilon_{p0}} = \varepsilon = \sin \omega t + 2k \sin (\omega t - \phi) \cos (\beta d \sin \theta)$$

$$= \sin \omega t + 2k \sin \omega t \cos (\beta d \sin \theta) \cos \phi - 2k \cos \omega t \cos (\beta d \sin \theta) \sin \phi$$

If we set $z = \beta d \sin \theta$,

$$\varepsilon = \sin \omega t (1 + 2k \cos \phi \cos z) - \cos \omega t (2k \sin \phi \cos z)$$

and

$$\begin{aligned} |\varepsilon| &= \sqrt{(1 + 2k \cos \phi \cos z)^2 + (2k \sin \phi \cos z)^2} \\ &= \sqrt{1 + 4k^2 \cos^2 z + 4k \cos \phi \cos z} \end{aligned}$$

since $k \ll 1$, we may disregard terms in k^2 , and

$$|\varepsilon| \approx \sqrt{1 + 4k \cos \phi \cos z} \quad (7-1)$$

To find maximum and minimum values of $|\varepsilon|$, we may take the derivative of the square of Eq. (7-1) and equate it to zero. Then

$$\frac{d|\varepsilon|^2}{dz} = 0 = -4k \sin z \cos \phi$$

This is satisfied for $\sin z = 0$ if $\cos \phi \neq 0$; and $\sin z = 0$ when $z = \beta d \sin \theta = n\pi$ and n is an integer. Then,

$$\sin \theta = \frac{n\pi}{\beta d} = \frac{n\lambda}{2d} \quad (7-2)$$

Substituting these values of z in Eq. (7-1), we obtain

$$|\varepsilon| = \sqrt{1 + 4k \cos \phi} \quad \text{for } n \text{ even; a maximum for } \cos \phi > 0 \quad (7-3a)$$

$$|\varepsilon| = \sqrt{1 - 4k \cos \phi} \quad \text{for } n \text{ odd; a minimum for } \cos \phi > 0 \quad (7-3b)$$

Thus the *positions* of the maximums and minimums are independent of the magnitude of the amplitude constant k and of the phase angle ϕ of the image radiators except insofar as the sign of $\cos \phi$ determines whether a given value of θ corresponds to a maximum or minimum.

TABLE 7-1

Width of ground screen, wavelengths	Minimums, deg		Maximums, deg	
	Measured	Calculated	Measured	Calculated
0.5	None	None	0	0
2.75	22	21.4	48	46.7
5.3	20	21	30	32
	44	46	64	66

Figure 7-6 shows a number of experimental patterns that were taken in the plane of polarization for slots cut in ground planes of varying size. Table 7-1 gives a comparison between the positions of the maximums and minimums calculated by formula (7-2), and those which were measured, in terms of angular displacement from the normal to the ground screen.

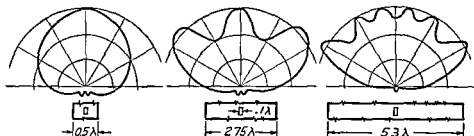


FIG. 7-6.—Measured radiation patterns in plane of polarization for waveguide-fed slot, showing variation in shape of pattern with size of ground plane.

From this analysis, it is evident that an irregular radiation pattern will result if a slot antenna is used in a small flat ground plane. As the width of the ground plane is increased, the number of maximums and minimums increases, but because of the decrease in k the magnitude of the ripple decreases. In aircraft installations, a ground plane that is large compared with the wavelength is generally presented, and a fairly

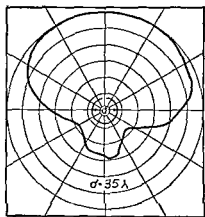


FIG. 7-7.—Measured radiation pattern in plane of polarization for resonant half-wave slot in cylindrical ground plane.

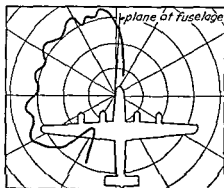


FIG. 7-8.—Measured radiation pattern in plane of polarization for vertical slot in side of airplane, showing curvature around nose of airplane.

smooth broad pattern will result. Edge effects may also be eliminated by the use of a ground plane that is curved at the edges or is cylindrical in shape. In the former case, the reflection at the edge of the ground plane is minimized; in the latter case, there is no edge to produce a reflection. Since the normal component of the electric field in the plane of

polarization exists at the surface of the ground screen, we should also expect some curvature of the wave front if the ground plane is curved in the plane of polarization. This is substantially verified experimentally, as shown in Figs. 7-7 and 7-8.

The effects of large obstacles encountered in aircraft installations of slot antennas, such as the wings, fuselage, or motors of the airplane, can generally be determined from line-of-sight considerations, since the curvature of the wave front around other than small obstacles is not very extensive. An example of this is shown in Fig. 7-9, which gives measured radiation pattern for a slot antenna mounted in the leading edge of the wing of a large four-engined airplane. In this case, the intensity of the measured field drops off rapidly in the region where the fuselage of the plane is interposed between the slot and the measuring system.

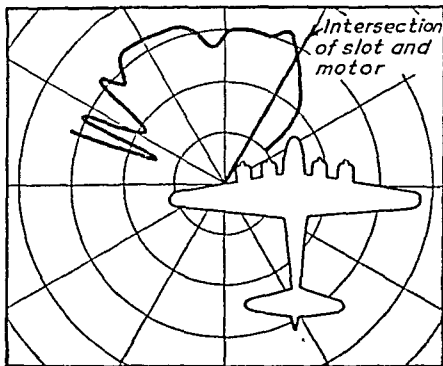


FIG. 7-9.—Measured radiation pattern in plane of polarization for vertical slot in leading edge of wing of airplane.

7-4. Effect of Width of Slot on Radiation Patterns.—Thus far, the discussion has been limited to the radiation patterns of slots that are essentially $\frac{1}{2}$ wavelength in width. If the slot is wider than $\frac{1}{2}$ wavelength, the equivalence between slot and dipole patterns still permits determination of the distribution of the radiation field, when the aperture is narrow. This fact is of particular importance in waveguide-fed slot antennas, since to satisfy cutoff requirements the guide must be somewhat greater than $\frac{1}{2}$ wavelength in width.

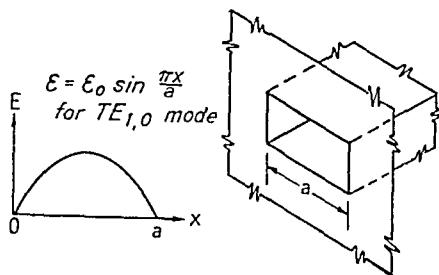


FIG. 7-10.—Electric field distribution at aperture of waveguide-fed slot, for $TE_{1,0}$ mode in guide.

in the Radio Research Laboratory. For practical use, we are concerned with the radiation patterns that obtain with a wave traveling down the guide and through the aperture in the fundamental or $TE_{1,0}$ mode. For this mode, the electric-field distribution at the aperture of width a will be of the form $E = E_0 \sin \pi x/a$, as shown in Fig. 7-10. The electric field will

The type of waveguide-fed slot antenna shown in Fig. 7-2a, consisting of a waveguide terminated directly in a ground screen, forms the basis of the broad-band slot antennas that have been constructed

have a single maximum value at the center of the slot until the frequency increases to such a value that the $TE_{2,0}$ and $TE_{3,0}$ modes may become dominant over the $TE_{1,0}$ mode. Because of the high cutoff frequencies of these modes, this distribution will be maintained at least until the width of the guide is somewhat greater than a full wavelength. Because of the equivalence of the electrical component of the radiation field of the slot and the magnetic component of the radiation field of a dipole, the radiation patterns produced by a narrow waveguide-fed slot antenna will be equivalent to those produced by a dipole of length equal to the width of the slot, provided that the current distribution on the dipole is equivalent to the electric-field distribution across the aperture of the slot. For simplicity, the width of the dipole is assumed to be negligible.

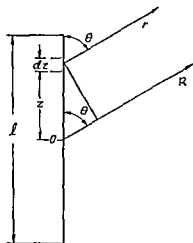


FIG. 7-11.—Dipole.

To compute the pattern of such a dipole, we may consider the radiation field produced by each of the infinitesimal dipoles comprising the length of the antenna. Referring to Fig. 7-11, we shall consider the field due to the elementary dipole of length dz which is at a distance z from the center of the dipole and at a distance r from the point at which the field is to be computed. Since this point is far removed from the dipole at a distance R from the center,

we may consider the vectors \vec{r} and \vec{R} to be parallel.

The electric component of the radiation field of the element dz is given by¹

$$d\mathcal{E}_\theta = \frac{-j\omega\mu_0 \sin \theta}{4\pi r} e^{j\beta r - j\omega t} u(z) dz \quad (7-4)$$

where ω = angular frequency, radians per second

$$j = \sqrt{-1}$$

μ_0 = permeability of free space = 1.257×10^{-6} henrys per m

β = phase shift constant = $2\pi/\lambda$

t = time, seconds

$u(z)$ = current at any point along the dipole

Upon integrating the contributions of all the elements along the length l of the dipole we obtain

$$\mathcal{E}_\theta = \frac{-j\omega\mu_0}{4\pi} e^{-j\omega t} \int_{-l/2}^{l/2} \frac{\sin \theta}{r} e^{j\beta r} u(z) dz \quad (7-5)$$

¹ See STRATTON, *op cit.*, p. 440.

If $r \gg l$ we may consider $\sin \theta/r$ invariant with respect to z and equal to $\sin \theta/R$. Then the integral may be rewritten as

$$\mathcal{E}_\theta = \frac{-j\omega\mu_0}{4\pi} e^{j\beta R - j\omega t} \frac{\sin \theta}{R} \int_{-l/2}^{l/2} e^{-i\beta z \cos \theta} u(z) dz \quad (7-6)$$

Since we wish to determine the field for a current distribution on the dipole equivalent to the voltage distribution at the aperture of the slot,

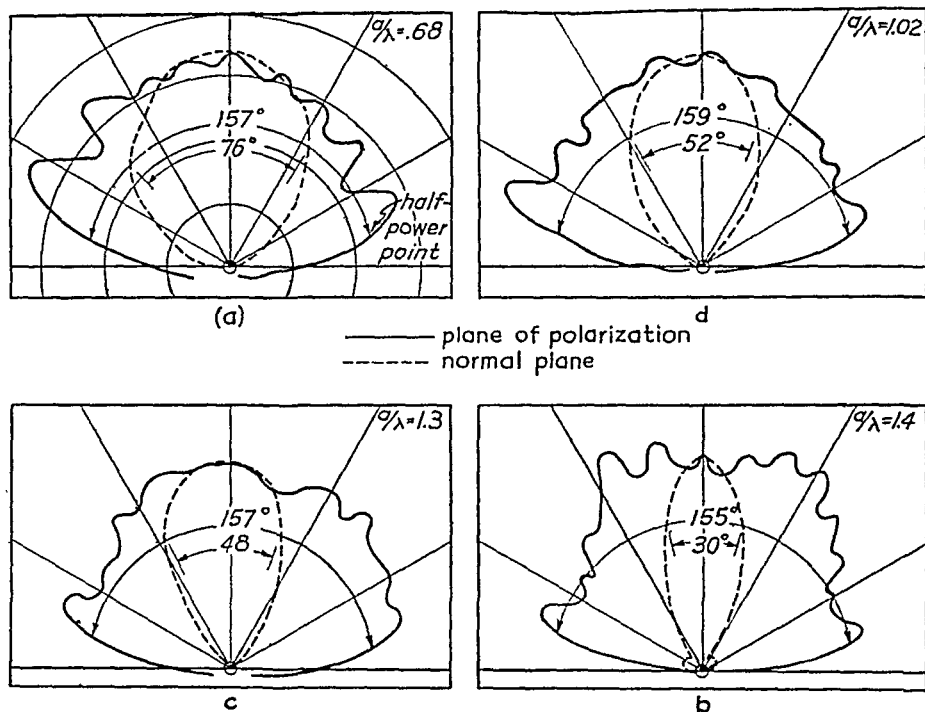


FIG. 7-12.—Measured radiation patterns of waveguide-fed slot of width a .

we may set $u(z) = I_0 \cos(\pi z/l)$. Substituting this value in Eq. (7-6) and solving the integral, we obtain a solution for \mathcal{E}_θ

$$\mathcal{E}_\theta = A l \left[\frac{\sin \theta \cos \left(\frac{\pi l}{\lambda} \cos \theta \right)}{1 - \frac{4l^2}{\lambda^2} \cos^2 \theta} \right] \quad (7-7)$$

where A is a constant.

For the case $l = \lambda/2$, Eq. (7-7) reduces to

$$\mathcal{E}_\theta = A' \left[\frac{\cos \left(\frac{\pi}{2} \cos \theta \right)}{\sin \theta} \right]$$

which is the usual expression for the field of a resonant half-wave dipole.

In the plane of polarization, the electric field given by Eq. 7-7, has a maximum value at right angles to the length of the dipole, the width of the pattern decreasing with increasing frequency. From the symmetry of the figure the field in the normal plane will have a constant value over 360 deg. Therefore, since the radiation pattern of a narrow¹ waveguide-fed slot antenna must be equal to one-half of these patterns with reversed polarization, the slot patterns should remain broad and symmetrical in the plane of polarization over a wide band of frequencies, limited only by the size of the ground plane. In the normal plane the field should have a single maximum value in the forward direction. Measured radiation patterns for a waveguide-fed slot are shown in Fig. 7-12.

7-5. Impedance Characteristics of Slot Antennas.—When a slot is fed with coaxial transmission line as in the example of Fig. 7-1, the impedance presented to the coaxial line is quite large compared with the values of characteristic impedance for common types of coaxial transmission line. Consequently, it is usually necessary to excite the slot at a point considerably removed from the center, as shown in Fig. 7-13. Such devices are inherently

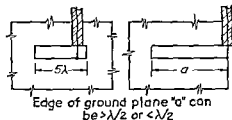


FIG. 7-13.—Coaxial-line-fed slots.

frequency sensitive, and coaxial-fed slots of these types are naturally narrow-band devices.

For this reason, it is desirable to consider other methods of exciting slots. The use of waveguide in feeding slots is indicated by the high center impedance of the slot, as well as by the physical aspects of the problem. The simplest type of waveguide-fed slot, shown in Fig. 7-2a, would appear to be inherently a broad-band structure, since a minimum discontinuity is involved in the path of the waves traveling through the aperture from the guide. No theory has yet been evolved to give the actual value of the impedance mismatch encountered in such a system, but these data are easily obtainable from waveguide measurements.

Figure 7-14 shows a number of standing-wave ratio curves that were taken for sections of waveguide of various width-to-height ratios terminated directly in a ground plane. It is evident from inspection of the figure that for guides with a width-to-height ratio less than about 3:1

¹ Experimental evidence indicates that a slot less than $\frac{1}{4}$ wavelength in height may, for practical purposes, be considered "narrow" in the sense that the pattern is equivalent to that of a dipole of negligible width. For slots of greater height, the pattern in the E plane will be somewhat narrower than 180 deg, although the pattern in the H plane will be unaffected.

the voltage standing-wave ratio introduced into the guide by the wave reflected at the aperture remains below 2:1 over a wide range of frequencies. The band is limited at the low-frequency end by the cutoff frequency of the guide. Since a standing-wave ratio of 2:1 is acceptable

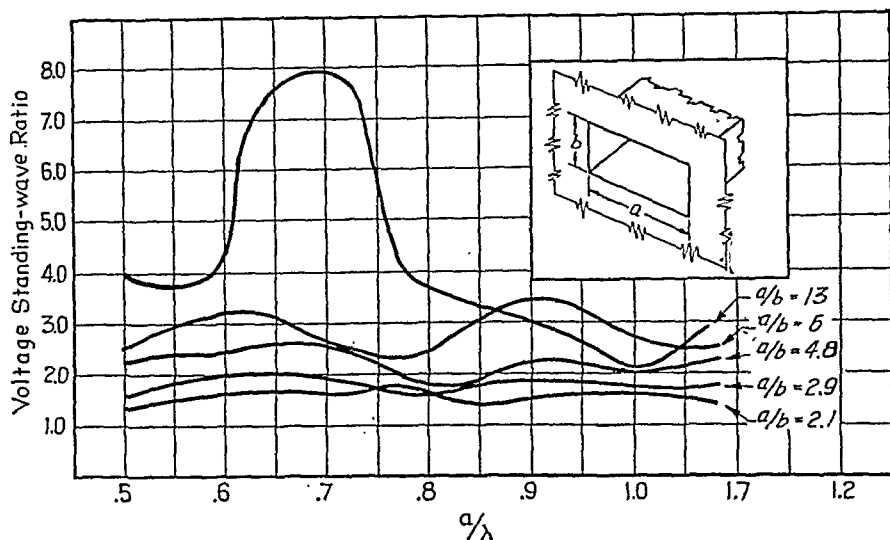


FIG. 7-14.—Standing-wave-ratio characteristics of waveguide terminated in ground screen.

for many purposes, it appears that a broad-band slot antenna could be constructed simply by terminating a section of waveguide directly in a ground plane. To be of practical use, however, such an antenna would have to include a broad-band coaxial-line-to-waveguide junction, so that the antenna could have a coaxial input. This junction could be designed to be as reflectionless as possible. If a completely reflectionless transformation were achieved, then the standing-wave ratio in the coaxial line would be the same as that in the guide and the antenna would be broad band. Transformers that approach this type of performance are in general fairly bulky and therefore not desirable. An alternative approach, and the one described below, consists of using a type of junction that not only is compact, but introduces reflections that tend to compensate those already present in the guide.

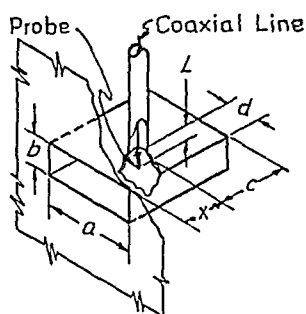


FIG. 7-15.—Waveguide-fed slot.

Let us consider the waveguide-fed slot of Fig. 7-15. This consists of a length of waveguide, short-circuited at one end, with the open end terminated in a ground screen. The cavity is excited by means of a

probe. The various parameters involved are

1. The width a of the waveguide
2. The height b of the waveguide
3. The distance c from the probe to the back plate of the cavity
4. The distance d from the probe to the nearer side of the cavity
5. The distance x from the probe to the aperture
6. The length L of the probe

We shall now analyze this waveguide-fed slot in an approximate way, to ascertain the orders of magnitude of the values that may be assigned to the various parameters to maintain a voltage-standing-wave ratio of less than 2:1 in a 50-ohm coaxial line connected to the probe, over a broad frequency band. The results of this analysis have checked experimental results very well in some cases. However, further work would be required to determine in general the validity of the results.

A waveguide is similar in many respects to a conventional transmission line.¹ We may consider energy radiated from a horn or aperture, or dissipated in a resistive medium in the guide, to be transferred from the guide to a lumped load impedance Z_L . We may also consider the values

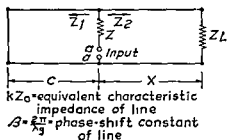


FIG. 7-16.—Equivalent transmission-line circuit for stub in waveguide cavity.

of voltage and current induced by the waves traveling within the guide and define a characteristic impedance Z_0 for the guide. For a section of air-filled rectangular waveguide in which energy is propagated in the $TE_{1,0}$ mode, Z_0 is given by²

$$Z_0 = \frac{\pi}{2} 377 \frac{b\lambda_g}{a\lambda} \quad \text{ohms} \quad (7-8)$$

where a and b are the width and height of the waveguide, respectively, λ is the free-space wavelength, and λ_g is the wavelength in the guide given by

$$\lambda_g = \frac{\lambda}{\sqrt{1 - \left(\frac{\lambda}{2a}\right)^2}}$$

for the $TE_{1,0}$ mode in air-filled rectangular guide.

¹ See Sec. 2-16, for example.

² The derivation of this formula is given in many texts. This is the value of characteristic impedance obtained by taking the ratio of the voltage between the top and bottom surfaces of the guide to the current flowing in the walls of the guide. This value of Z_0 differs by a constant factor from that obtained by taking the ratio of the power flowing through the guide to the square of the current.

In our analysis, we may replace the waveguide components with equivalent transmission-line components, as in Fig. 7-16. In this figure, the waveguide, of characteristic impedance Z_0 , is replaced by a parallel-wire transmission line of characteristic impedance kZ_0 , and the load impedance Z_L is similarly replaced by kZ_L , where k is a constant factor that depends on the length of the probe and its relative position with respect to the side of the guide and is given by¹

$$k = \frac{2L^2}{b^2} \sin^2 \left(\frac{\pi d}{a} \right) \quad (7-9)$$

where a , b , d , and L are the dimensions shown in Fig. 7-15.

The impedance seen at the input terminals aa is equal to the sum of a series impedance Z and an impedance Z_p , which is the parallel combina-

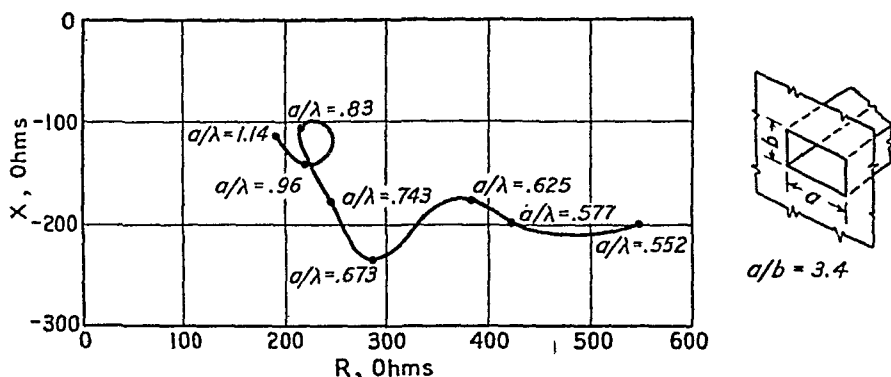


FIG. 7-17.—Impedance characteristic at aperture of waveguide terminated in ground plane.

tion of the load impedance translated to the feed point (Z_2) and the impedance of the short-circuited section of guide forming the back of the cavity (Z_1). The series impedance Z is the intrinsic impedance of the stub, due to its electrical length, diameter, and conductivity, plus the impedance of any matching sections placed in series with the stub, either between the input terminals and the stub, or between the stub and the guide. This analysis neglects second-order effects that may be produced by the proximity of the stub to the aperture and back plate.

The load impedance Z_L can best be determined experimentally by terminating a section of waveguide directly in a ground plane and taking impedance measurements by means of a movable probe in the guide. An experimental curve giving the results of such measurements for a section of waveguide with a width-to-height ratio of 3.4:1 is shown in Fig. 7-17. It is noted that the reactance at the aperture is capacitive over the entire frequency range. The impedance introduced at the feed

¹ SLATER, J. C., "Microwave Transmission," Chap. VII, McGraw-Hill Book Company, Inc., New York, 1942.

point by the short-circuited back portion of the waveguide cavity is, neglecting attenuation in the surface of the guide, a pure reactance $X_r = jZ_0 \tan 2\pi c/\lambda_g$. The variable c is the distance between the feed point and the back plate. If $c < \lambda_g/4$, this reactance will be inductive. If the distance x , between the feed point and the aperture, is also small, the reactance X_r will tend to compensate the capacitive reactance at the aperture. In the circuit of Fig. 7-16, $Z_1 = kX_r$.

Figure 7-18 shows the computed resultant impedance at the feed point for a cavity with relative dimensions $b/a = 0.295$, $c/a = 0.2$, $x/a = 0.125$. The numerical values of resistance and reactance are given in terms of the characteristic impedance Z_0 of the guide computed

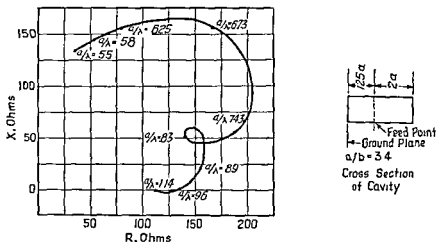


Fig. 7-18.—Impedance characteristic at feed point of short-circuited cavity.

by formula (7-8). In the equivalent transmission-line circuit, the impedance of Fig. 7-18 is equal to Z_r/k .

The factor k may be given any value between 0 and 2 by variation of the length of the stub and the distance from the center of the stub to the side wall of the cavity. The intrinsic impedance of the stub may be kept low by making the diameter of the stub large. We may also compensate for the large inductive reactance introduced by the short-circuited back portion of the cavity at the low-frequency end of the band by means of a series capacitive reactance between the stub and the walls of the guide. The net impedance of the stub, Z_1 will then be a capacitive reactance. One method of accomplishing this is by connecting the stub at its end to a long conductor, the ends of which are connected to the side walls of the cavity, as in Fig. 7-19. This method of matching has the advantage that it is internal to the structure and does not increase the total size of the antenna.

A bar-matching section, such as illustrated in Fig. 7-19, may be

considered a section of coaxial transmission line, short-circuited at each end, the bar forming the inner conductor and the walls of the cavity forming the outer conductor. With the dimensions given, the characteristic impedance of this transmission line can be computed to be about 66 ohms. Since the stub is centered in the line, the reactance seen at the base of the stub due to the compensating bar is approximately $X_s = 6\frac{1}{2} \tan \pi a/\lambda$. If a is greater than $\frac{1}{2}$ wavelength, as it must be to satisfy cutoff requirements, this reactance will be capacitive for $a/\lambda < 1$.

The insertion of a bar into the guide adds a lumped parallel reactance X_b at the point. This is equivalent to placing a reactance kX_b in

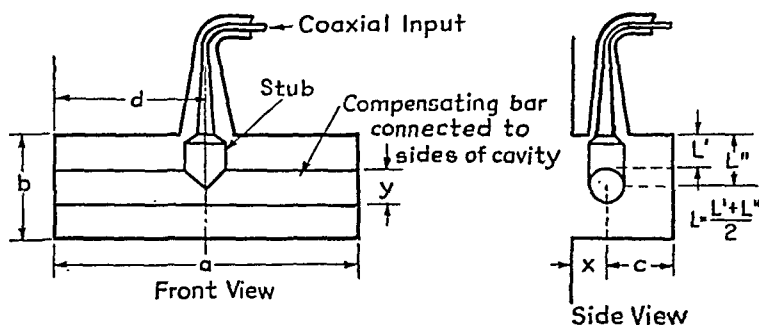


FIG. 7-19.—Waveguide-fed slot antenna with compensating bar.

parallel with Z_1 and Z_2 in the equivalent transmission-line circuit of Fig. 7-16. The value of X_b can be determined by the following formula¹ (refer to Fig. 7-19 for various parameters):

$$\frac{Z_0}{X_b} = \frac{-8b}{\lambda_g} \left\{ \log \csc \pi \frac{(b-y)}{2b} + \left[\frac{1}{\sqrt{1 - \left(\frac{b}{\lambda_g}\right)^2}} - 1 \right] \frac{\cos^4 \pi \frac{(b-y)}{2b}}{1 + \left[\frac{1}{\sqrt{1 - \left(\frac{b}{\lambda_g}\right)^2}} - 1 \right]} \right. \\ \left. \sin^4 \pi \frac{(b-y)}{2b} + \frac{1}{16} \left(\frac{b}{\lambda_g}\right)^2 \left[1 - 4 \sin^2 \pi \frac{(b-y)}{2b} + 3 \sin^4 \pi \frac{(b-y)}{2b} \right]^2 \right\} \quad (7-10)$$

We may then define an impedance Z_p' which is given by

$$\frac{1}{Z_p'} = \frac{1}{Z_p} + \frac{1}{X_b}.$$

The values of Z_p'/k shown in Fig. 7-20 differ only slightly from the values of Z_p/k shown in Fig. 7-18.

¹ MARCUVITZ, N., "Waveguide Handbook," Radiation Laboratory Series, Vol. 10, McGraw-Hill Book Company, Inc., New York (in press).

To compute the value of k by Eq. (7-9) we must first determine the length of the stub. Referring to Fig. 7-19, we shall take

$$L = \frac{L' + L''}{2}$$

as an approximation. Then, since $d = a/2$ and $L = 0.411b$ we can compute $k = 0.34$.

From the circuit of Fig. 7-16, we know that the input impedance Z_i for the entire structure is given by

$$Z_i = Z + Z_p'$$

We may set $Z = X_s$, neglecting the intrinsic impedance of the stub. Values of Z_p' may be obtained from Fig. 7-20.

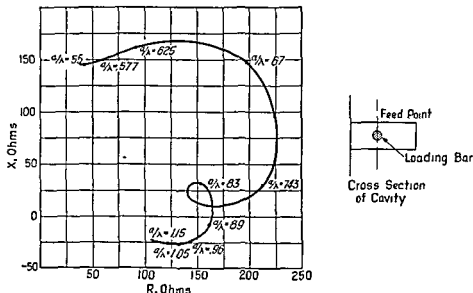


FIG. 7-20.—Impedance characteristic at feed point of short-circuited cavity with loading bar.

The resultant input impedance for the combination of parameters chosen is shown in Fig. 7-21. A circle of 2:1 standing-wave ratio on a 50-ohm line is drawn on this curve. It is evident that this system provides a match over a wide frequency range with this 2:1 standing-wave-ratio limit of acceptability. An experimental impedance curve for this type of antenna is shown in Fig. 7-22. The agreement between the theoretical and experimental data is fairly good.

For a guide-fed slot of given cross section, the important parameters then are the length of the stub, the distance from the aperture to the stub, the diameter of the loading bar, and the distance from the stub to

the back of the cavity. The values of the parameters chosen in the above analysis are certainly not unique, nor do they necessarily represent an optimum choice. Experimentally it has been found that slight variations in any of the parameters are not excessively critical, and the mechanical tolerances involved are simple.

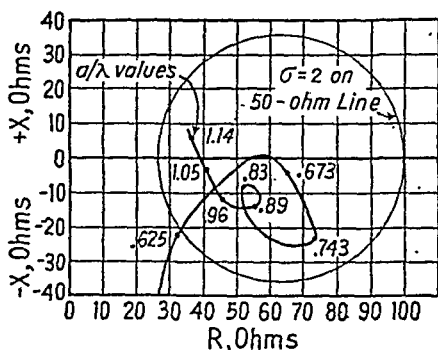


FIG. 7-21.—Theoretical impedance characteristic at input of compensated slot antenna of Fig. 7-19. $a/b = 3.4$.

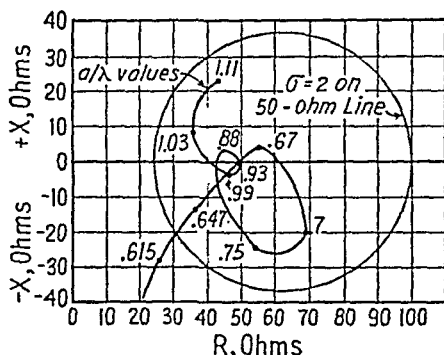


FIG. 7-22.—Experimental impedance characteristic at input of compensated slot antenna. $a/b = 3.2$.

7-6. Methods of Reducing Size of Antenna.—The rectangular waveguide-fed slot, while it has broad-band characteristics, has the disadvantage that it is large in comparison to the wavelength, being about $\frac{6}{10}$ wavelength wide at the low-frequency end of the band.

Some reduction in size can be achieved by employing sections of waveguide in which the cutoff wavelength is greater than half the width of the guide. This can be realized by the use of *capacitance-loaded* guide, or *open* guide.

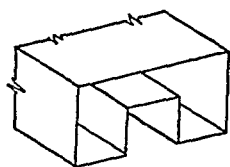


FIG. 7-23.—Capacitance-loaded waveguide.

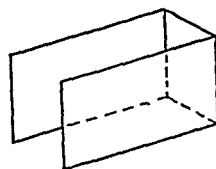


FIG. 7-24.—Open waveguide.

Capacitance loading can be attained by distorting a section of waveguide, or by the use of dielectric-filled guide. If the guide is distorted in the manner of Fig. 7-23, such that the top and bottom surfaces are closer together at the center of the guide than at the edges, it has a considerably lower value of characteristic impedance and cutoff frequency than a section of unloaded guide of comparable dimensions. The mismatch involved between the end of the guide and free space is fairly large, depending on the degree of loading. Therefore slot antennas employing sections of loaded waveguide are considerably more frequency sensitive than unloaded slots.

The use of sections of open waveguide, i.e., waveguide consisting only of two planes, is advantageous since cutoff considerations are removed entirely in such a system (Fig. 7-24). However, a slot antenna constructed with open guide is not restricted to radiating only in a forward direction. Therefore, this type of slot is usable only in cases where a severely curved protruding ground plane is presented. In practice, this may be realized by installation of a slot in the leading edge of the wing or wing tip of an airplane.

7-7. Examples of Slot Antennas.—The major emphasis in the slot work at the Radio Research Laboratory was on designing slot antennas with broad-band frequency characteristics for use in high-speed aircraft. Some additional design work was done on narrow-band slots and on slots in cylindrical ground planes. In this section we shall limit the discussion to those slots which were developed at this laboratory, giving frequency characteristics and design data for each type.

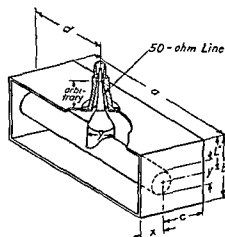


FIG 7-25 —T-fed slot antenna.

Figure 7-25 shows the compensated slot antenna of the type discussed in Sec. 7-5. Because of the shape of the feed section, this type of slot is referred to as a *T-fed* slot antenna. Dimensions of two T-fed slots of different width-to-height ratios are given in Table 7-2. These dimensions are in terms of the maximum wavelength at which the voltage standing-wave ratio introduced into a 50-ohm coaxial line connected to the input of the antenna falls below 2:1.

TABLE 7-2

Dimensions (from Fig 7-25)	Antenna I	Antenna II
a	0.62 λ	0.65 λ
b	0.20 λ	0.14 λ
c	0.127 λ	0.113 λ
d	0.31 λ	0.325 λ
y	0.073 λ	0.055 λ
x	0.073 λ	0.07 λ
L''	0.10 λ	0.07 λ

Voltage standing-wave ratio characteristics of these two antennas are shown in Fig. 7-26. If either of these slots is mounted in a large flat ground screen, the radiation patterns obtained are equivalent to those given in Fig. 7-12. Since

the long portion of the T-shaped feed section carries considerable current and is close to the aperture, it might be feared that some energy would be radiated in the form of a "butterfly" pattern with the electric vector parallel to the long dimension of the slot. Although measurements indicate that such radiation is present, the measured maximums are about 16 db or more below the maximums for the normal slot-type radiation.

Two slot antennas that employ sections of H-shaped capacitance-loaded waveguide are shown in Fig. 7-27. The H-shaped sections are used to reduce the over-all width of the slot for operation at a given frequency. The effective width of the guide is increased by distorting the structure in the manner of the figure, effecting essentially a capacitance loading at the center of the guide. This loading may be further increased by addition of dielectric material filling the entire guide as shown, or merely filling the narrow space between the top and bottom edges. The slot shown in Fig. 7-27a is energized by a T-shaped feed section, as in the previously described slot. The addition of a thin layer of dielectric around the long portion of the feed section increases its effective electric

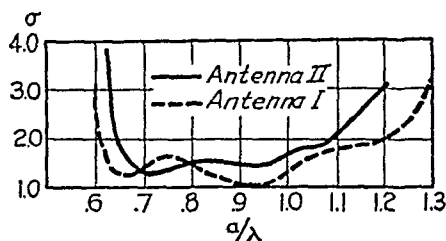


FIG. 7-26.—Standing-wave ratios on 50-ohm line of T-fed slot antennas.

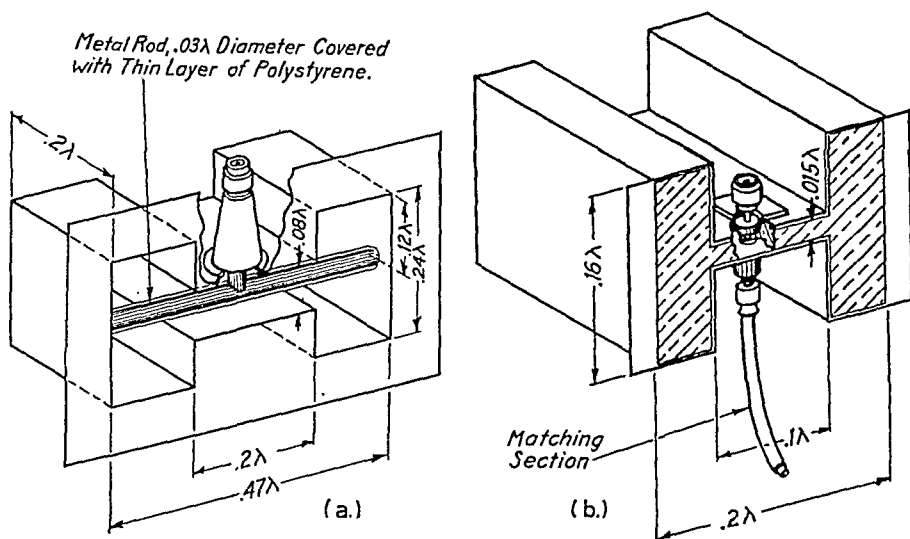


FIG. 7-27.—(a) Loaded T-fed slot antenna; (b) Dielectric-filled loaded slot antenna.

length in accordance with the increase in width of the loaded waveguide. The voltage standing-wave ratio introduced into a 50-ohm line connected to the input of the antenna remains below 5:1 over about a 2.5-to-1 frequency band. The standing-wave ratio falls below 2:1 in voltage over a narrow band at the low-frequency end of the range.

The slot of Fig. 7-27b demonstrates a case of extreme capacitance loading.

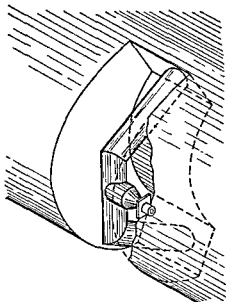


Fig. 7-28.—Curved T-fed slot antenna.

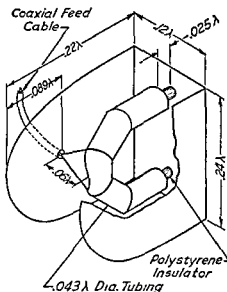


Fig. 7-29.—Curved slot antenna.

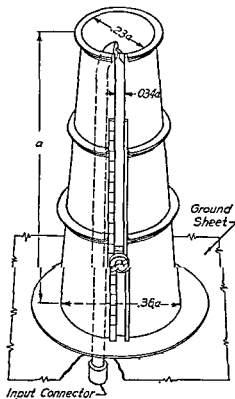


Fig. 7-30 —“Split-can” antenna.

This slot is very small compared with unloaded slots, and the frequency band is extremely narrow.

To satisfy certain installation requirements, particularly in aircraft, it is sometimes necessary to construct a slot to conform to a rather severe curvature. This type of slot is illustrated in Fig. 7-28. Here the front of the cavity is shaped to conform to the curvature of the ground screen. The long portion of the T feed section is also curved. Internal relative dimensions of this type of slot may be kept equivalent to those of the straight rectangular slot for comparable performance. It should be pointed out that in installations of this type the shape of the radiation pattern will be controlled to a large extent by the shape of the curved surface, and that in general these patterns will be very different from those obtained with slots in flat or almost flat surfaces.

Another type of curved slot that

is particularly applicable at low frequencies is shown in Fig. 7-29. This is relatively smaller in wavelengths than a corresponding curved slot modified from the straight rectangular type. Since the curved section of guide has degenerated into a sort of metal channel, waveguide considerations are no longer a limiting factor in this structure. The system may be excited by means of a stub that is capacitance loaded by a massive curved bar connected to the end, as shown, and

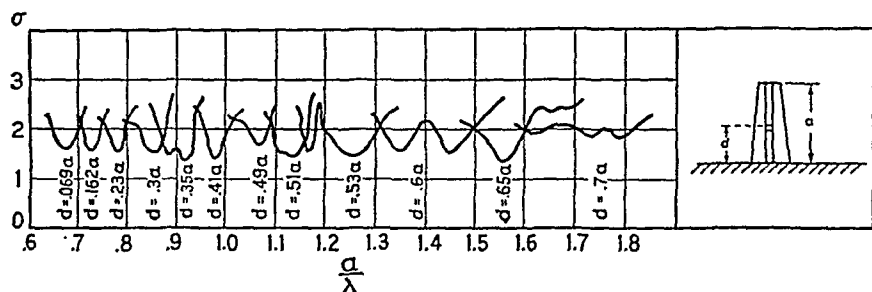


Fig. 7-31.—Standing-wave-ratio characteristics on 50-ohm line of "split-can" antenna.

introduces a standing-wave ratio of less than 2:1 on a 50-ohm line over about a 40 per cent frequency band. This type of slot is limited in use to an installation in which a severely curved protruding surface is used for mounting, such as the leading edge or tip of an airplane wing.

The slot-antenna structure shown in Fig. 7-30 is known as a *split-can* antenna. This antenna consists of a tapered ground plane in the form of a truncated cone, in which a long slot has been cut. The slot is excited by means of a coaxial line connected across one end. The length of the slot is adjustable by means of a sliding short-circuiting bar. For a given setting of the short-circuiting bar, the antenna is rather frequency sensitive (band width of 5 to 10 per cent) as would be expected with this direct coaxial feed. However, by proper adjustment of the slider for each frequency, the antenna can be operated over a considerable frequency range. The ground plane is tapered so that the mean diameter in wavelengths of the section of ground plane between the feed point and the short circuit will remain approximately constant with frequency, and the patterns will tend to remain uniform over a wide range. The standing-wave-ratio characteristics of this antenna, for various positions of the short-circuiting bar, are indicated in Fig. 7-31. Measured radiation patterns in the plane of polarization are shown in Fig. 7-32.

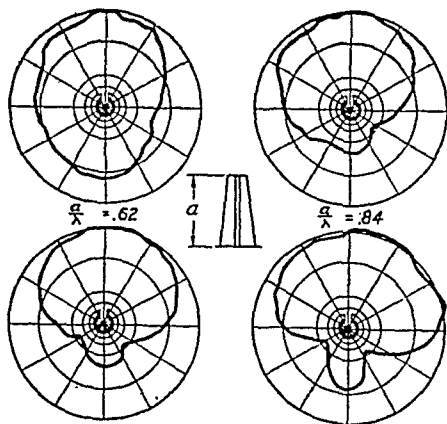


Fig. 7-32.—Radiation patterns in plane of polarization of "split-can" antenna.

CHAPTER 8

AUXILIARY DEVICES FOR ANTENNAS

By E. L. BOCK AND J. MARGOLIN

8-1. Broad-band Microwave Switches: General Considerations.—

Possible uses for microwave switches are both widespread and obvious. Any system that is designed to cover an extremely large frequency range or to perform several varied functions is likely to require switching equipment. For example, the frequency range of a particular transmitter (or receiver) may be greater than that obtainable with the antenna of

the type with which it is to be used. It then becomes necessary to use several antennas, each of which covers part of the frequency range. Again, in the case of a receiving system, it is often desirable to be able to choose between several antennas differing in polarization and directional properties. In such circumstances, a convenient switching mechanism is very useful.

The design problems associated with microwave switching mechanisms differ in character from the problems associated with most other microwave devices in that they originate primarily in connection with mechanical rather than electrical func-

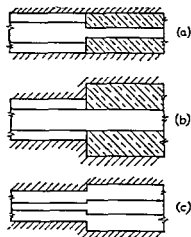


FIG. 8-1.—Types of discontinuities.

tions. Obviously, the electrical problem consists of keeping reflections small. In accordance with general broad-band design principles the best means of achieving this is to keep individual reflections small rather than to permit large reflections and attempt to cause them to cancel each other.

All the devices described in this chapter, with the exception of the waveguide switch of Sec. 8-4, are designed for use with coaxial cable. Figure 8-1 shows a few of the types of discontinuities that are frequently necessary in such coaxial devices to achieve satisfactory mechanical designs. Since each of these examples represents a junction between two coaxial lines of the same characteristic impedance, the reflections that they introduce cannot be computed by means of ordinary transmission-line theory. A method for finding the approximate values of the reflec-

tion coefficients to be expected from discontinuities of these types has been obtained by Whinnery, Jamieson, and Robbins.¹ They have shown that the effect of such a discontinuity is equivalent to that which would be produced by a lumped capacitance shunted across a uniform transmission line at the junction and have devised procedures for evaluating the magnitude of this capacitance.

Since the magnitude of the capacitance is determined only by physical dimensions and is not dependent on frequency, the magnitude of the shunt admittance, and therefore the magnitude of the reflections, will increase continuously with increase of frequency. This is illustrated by the curve of Fig. 8-2, which shows the computed standing-wave ratio that a discontinuity like that of

Fig. 8-1c would produce in an otherwise reflectionless line. This example has been chosen because it is representative of the discontinuities occurring in devices intended for use with standard flexible solid-dielectric cable. The dimensions used in the computations are almost exactly those which occur in a commonly used connector for flexible 50-ohm cable. A drawing of this connector is shown in Fig. 8-3. It should be noted that the standing-wave ratio begins to be appreciable at commonly used frequencies. For example, at 3500 Mc it is 1.1:1. If the dimensions of

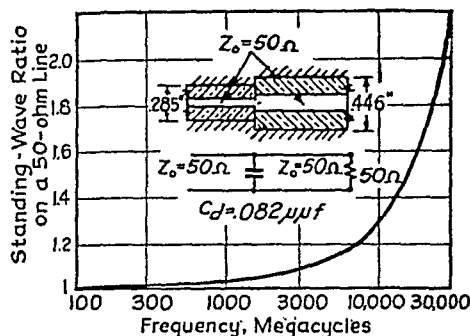


FIG. 8-2.—Computed standing-wave ratio produced by a typical discontinuity in coaxial cable.

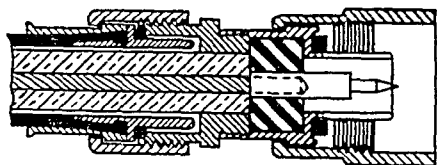


FIG. 8-3.—Type-N connector for 50-ohm flexible cable.

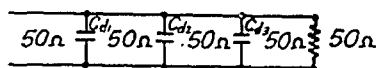
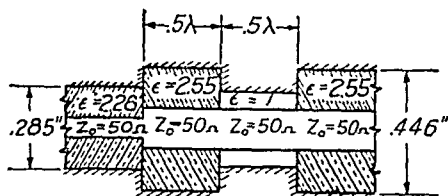


FIG. 8-4.—Example of multiple discontinuities: (a) three discontinuities in 50-ohm line; (b) equivalent circuit.

the cable (and all other dimensions) were smaller, this standing-wave ratio would occur at a proportionally higher frequency.

In any actual switch or connector, several abrupt junctions are virtually inevitable. Since the general behavior of systems involving mul-

¹ WHINNERY, J. R., JAMIESON, H. W., and T. E., ROBBINS, Coaxial Line Discontinuities, *Proc. I.R.E.*, **32**, 695 (1944).

multiple discontinuities has been discussed in Sec. 1-5, all that is presented here is the quantitative example of Fig. 8-4 which shows three discontinuities like the one referred to above. The standing-wave ratio that this combination produces on an otherwise flat line at 3000 Mc is 1.22:1. This is the case where the discontinuities are $\frac{1}{2}$ wavelength apart, the worst case, as has been shown in Sec. 1-5. Since in any combination of connectors and switches there will normally be more than three discontinuities, it can be seen that the standing-wave ratio of the combination will vary considerably with frequency, and even though the individual discontinuities are small, the resultant is likely to be quite large at some frequencies. In this connection it may be noted that the measurements of standing-wave ratio of all coaxial switches were made with the connectors normally used with them present.

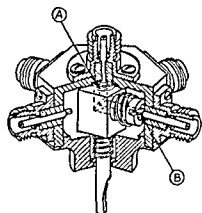


FIG. 8-5.—Single-pole, six-position coaxial switch.

perform the switching operation. Two of these are described in this section. The single-pole six-position switch shown in Fig. 8-5 is intended primarily for receiving applications up to 3500 Mc. Its insertion loss in the range of 3500 to 10,000 Mc, however, is less than 6 db, and if this loss can be tolerated the switch may be used up to the higher limit. The measured standing-wave ratios produced by a production model of this switch on a 50-ohm transmission line are plotted in Fig. 8-6, and it will be seen that the switch introduces very small reflections at frequencies up to 3000 Mc.

It can be seen from the drawing that the junction at A in Fig. 8-5 is made by butting the outer conductor of the rotating section of coaxial line against a bearing surface on the switch housing. A compression spring around the knob shaft keeps the outer conductor of the rotor in contact with the bearing surface. The center conductor is rigidly supported in the dielectric-filled rotating section of coaxial line and is drilled undersize and slotted to form a tight-fitting joint with the cylindrical center conductor of the common connector.

8-2. Coaxial Radio-frequency Switches.—There are numerous possible types of switches in which a short section of coaxial line is rotated to

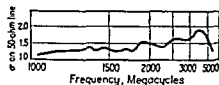


FIG. 8-6.—Standing-wave ratio produced by six-position coaxial switch.

As the radial section of coaxial line rotates about the knob shaft it contacts each of the radial chassis connectors. These connectors are arranged with their center conductors flattened and extended into the switch housing. The center conductor of the rotating section is slotted and compressed (at *B* in Fig. 8-5) to grip the flattened center conductors of the radial connectors. Outer conductor contact is made by a coin-silver sleeve sliding on a set of fingers at the end of the rotating section of coaxial line. The sleeve is held against the inner wall of the switch housing by a compression spring and is slotted to allow it to clear the protruding center conductors of the radial connectors.

The magnitude of the reflections produced by a switch of this type will depend greatly not only on the accuracy of fabrication, but also on the accuracy with which the rotor is centered at each position. In the manually operated versions, sufficiently accurate centering has been accomplished by a simple detent mechanism consisting of a spring-driven steel ball that is forced into a spherical depression in the rotor assembly when the rotor is in position.

This switch has been provided with several types of remote driving mechanisms. One of these is the Geneva movement¹ which can be

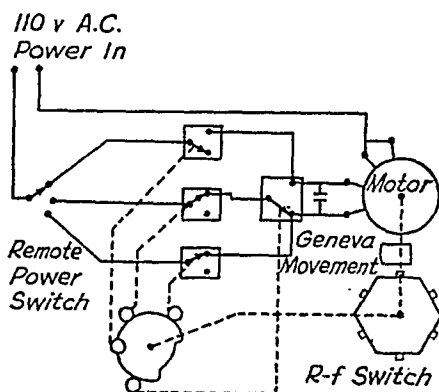


FIG. 8-7.—Three-position actuator for six-position coaxial switch.

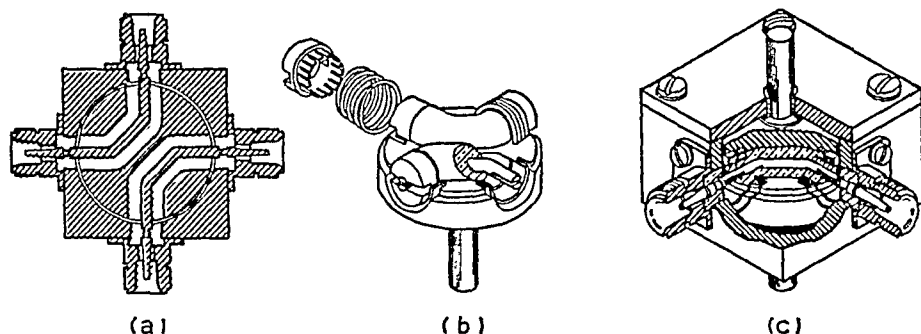


FIG. 8-8.—Crossover switch.

machined to provide very accurate positioning. Controlled by a remote multiposition power switch and suitable contactors on a cam on the shaft of the r-f switch, a motor can be made to drive the Geneva movement. The diagram of a typical three-position remote actuator is shown in Fig. 8-7.

¹ *J. Soc. Motion Picture Eng.*, 44, 245 (1945).

The crossover switch of Fig. 8-8 is also a coaxial switch, and the contacts are similar to those at the radial connectors of the six-position switch described above. As can be seen in Fig. 8-8, 90-deg rotation of the rotor of this switch will interchange the connections. This switch can be actuated either manually or remotely as described above. The measured standing-wave ratios produced by the crossover switch on a 50-ohm transmission line in a small frequency range are plotted in Fig. 8-9.

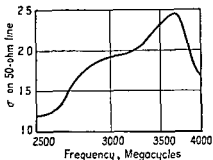


FIG. 8-9—Standing-wave ratio of crossover switch.

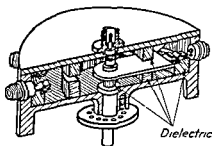


FIG. 8-10.—Twelve-position flat-bladed switch

Although the data presented for both this and the six-position switch were measured only in limited frequency ranges, it may be assumed that the standing-wave ratios that they produce become progressively lower with decreasing frequency.

8-3. Flat-bladed Switches and Relays.—A simple r-f switch for use with coaxial transmission lines of 50-ohm characteristic impedance is shown in Fig. 8-10. The rotating radial arm of this 12-position switch is simply a flat blade proportioned in such a manner that the transmission

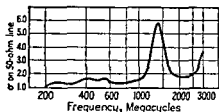


FIG. 8-11.—Standing-wave ratio produced by flat-bladed switch.

line formed by the top and bottom walls of the cylindrical housing and the blade has a characteristic impedance of approximately 50 ohms. The discontinuities at the ends of the blade are considerably larger than those in the switches of Sec. 8-2, but at lower frequencies, where the dimensions are small in terms of wave-

lengths, the reflections produced are quite small. The curve of measured standing-wave ratios is shown in Fig. 8-11. Though flat-bladed switches will operate satisfactorily only at the lower end of the microwave region, the simplicity of construction often recommends their use.

Figure 10-35 shows a single-pole double-throw switch that has been used at frequencies up to 5000 Mc. It consists of three sections of flat-bladed transmission line (characteristic impedance approximately 50

ohms) in the form of a Y. The outer conductors of the three lines are channels of rectangular cross section milled in a split metal block. The center conductor of the transmission line that forms the vertical leg of the Y extends into the small cavity at the junction of the three lines. The center conductors of the transmission lines forming the two diagonal arms of the Y are flat flexible leaves that also extend into the cavity, one above and one below the rectangular conductor. Two formica rods, one coupled to each leaf, extend through holes in the relay housing to the armature of a solenoid. Depending on the position of this armature, one or the other of the leaves is moved into contact with the common conductor. To minimize cross talk between the common line and the unused line, the center conductor of the unused line is short-circuited to the switch housing.

This type of switch has been used to select remotely the vertical or horizontal element of the direction-finder antennas described in Secs. 10-26 to 10-34.

8-4. Waveguide Switch.—Transmission circuits for the centimeter range of radio frequencies usually make use of waveguide. Since wave-

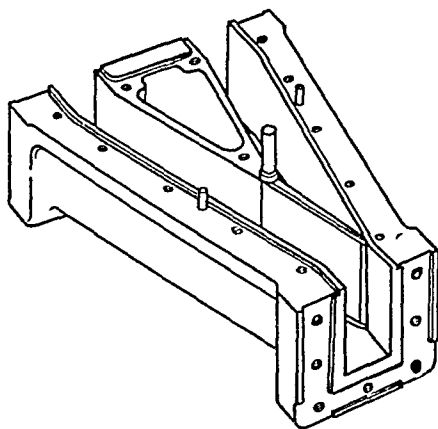


Fig. 8-12.—Waveguide switch (cover removed).

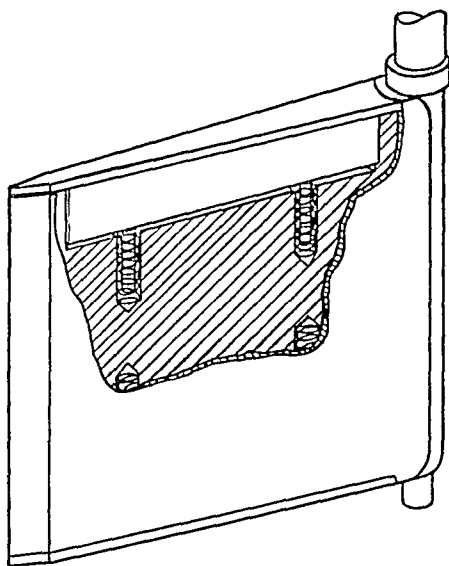


Fig. 8-13.—Vane for waveguide switch.

guide connections are not in general easily changed, switches are essential if changes are to be made at all frequently. For rectangular waveguide the single-pole double-throw switch of Fig. 8-12 has been designed. The switching is achieved by means of a gate or vane pivoted to block off one arm of the Y-junction of three waveguides.

Contact between the top and bottom walls of the guide and the vane

is made by means of the sliding contactors shown in the drawing of Fig. 8-13. The compression springs expand the vane assembly so that the unused section of guide is completely blocked off by it. Small amounts of wear are taken up by the springs. The standing-wave ratio introduced by a switch of this type designed for 10-cm waveguide¹ is shown in Fig. 8-14. Leakage into the unused section of guide was determined to

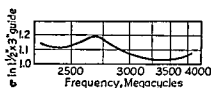


FIG. 8-14.—Standing-wave ratio produced by waveguide Y switch.

be more than 60 db down, and an experimental model showed no signs of arcing with approximately 1 kw of power passing through the switch.

A method for actuating the switch is shown in the diagram of Fig. 8-15. A lever arm attached to the shaft of the vane is driven through a compression spring link by an idler gear that is coupled to a reversible

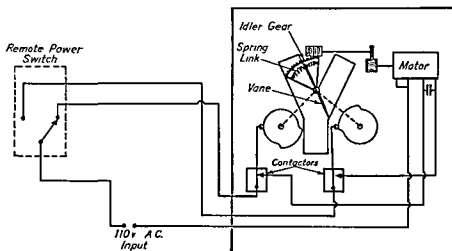


FIG. 8-15.—Waveguide-switch actuator.

motor by suitable worms and worm gears. Once energized, the motor drives the vane to position, a contactor in series with the motor power circuit is opened by a cam, and the overdrive of the motor is taken up by the compression-spring link. The vane is held in position by the spring. A simple lever arm for the vane shaft can, of course, be hand-operated.

¹ Waveguide 3 by $1\frac{1}{2}$ in. by 0.085-in. wall. λ cutoff approximately 14.5 cm.

CHAPTER 9

PRINCIPLES OF DIRECTION FINDING

BY A. ALFORD, J. D. KRAUS, AND E. C. BARKOFSKY

9-1. The General Direction-finder Problem.—A radio direction finder may be defined as a device for determining the direction of arrival of r-f energy. It is a receiving system and operates on the energy that it extracts from passing radio waves.

The direction of flow of energy from a radio source or transmitter in free space is along radial lines from the transmitter. The basic direction-finder problem is to determine the direction in which r-f energy is received from a distant source. Thus, a fundamental requirement of a direction-finder antenna is that it possess a property by which the direction of arrival of r-f energy can be determined.

9-2. Wave Polarization and Cross Field. Plane Polarization.—If the direction of the electric field of an electromagnetic wave is confined to a single plane, the wave is spoken of as *plane polarized*. The plane determined by the direction of the electric field and the direction of energy flow is called the *plane of polarization*, or the *plane of the electric vector*. When the plane of the electric vector is parallel to the surface of the earth, the wave is said to be *horizontally polarized*.

Elliptical and Circular Polarization.—The electric-field vector \mathcal{E} is shown at a point A in space in Fig. 9-1a. The direction of propagation is perpendicular to the page and toward the reader. The angular position of the electric vector with



FIG. 9-1.—(a) Relation of magnitude of electric-field vector \mathcal{E} and angular position θ . (b) When the angle θ is constant the wave is plane polarized.

respect to a reference plane is given by the angle θ . If θ is constant, the electric vector is confined to a single plane, as shown in Fig. 9-1b, and the wave is plane polarized. The values of \mathcal{E} shown are the maximum positive and negative values.

If the magnitude of the electric vector \mathcal{E} at point A remains constant while the angle θ increases uniformly with time so that it changes by 360 deg in one period, the wave is *circularly polarized*. The positions of the electric vector at intervals over a complete cycle are shown in Fig. 9-2a for a wave traveling in a direction perpendicular to the page.

If the angle θ changes by 360 deg during one cycle, but the magnitude of the electric vector changes over the cycle, the wave is *elliptically polarized*. The direction and magnitude of the electric vector of an elliptically polarized wave at various instants over a complete cycle are shown in Fig. 9-2b for the case where the maximum field is vertical, in Fig. 9-2c for the maximum at 45 deg.

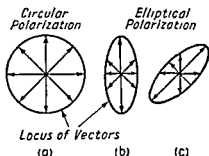


FIG. 9-2.—Direction of electric field in plane perpendicular to direction of propagation for circularly and elliptically polarized waves. Arrows indicate direction of field at a point in space at various instants of time over a complete cycle.

the rate of rotation of the electric field is 360 deg per period. If the electric vector rotates at a rate different from this, the wave may be considered to have *rotating-plane polarization*. For example, a half-wave element can be rotated around an axis perpendicular to itself by means of a motor. If the antenna is supplied with r-f energy via a transmission line with rotating joint, a plane-polarized wave will be radiated but the plane of polarization will rotate at a rate determined by the speed of the motor.

Elliptical Cross Field.—There is another type of field condition that can be referred to as *cross field*. In general, cross field involves the variation of the electric vector in both magnitude and direction in the plane of polarization. This type of field may result when waves from two plane-polarized sources of the same frequency arrive at a point from directions making an angle ϕ in the plane of polarization. Such a field is illustrated in Fig. 9-3, in which S_1 and S_2 are two sources of the same frequency polarized in the plane of the page.

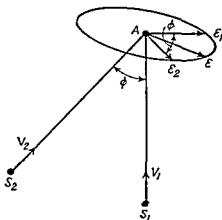


FIG. 9-3—Elliptical cross field resulting from two waves of same frequency polarized in the plane of the page, arriving from different directions. In this example the time-phase difference between the two waves at point A is taken as 90 deg.

The two sources may be two antennas connected to the same transmitter. Or one may be a primary source and the other a reflecting object or secondary source. This latter combination is discussed in more detail in Sec. 9-7. Arrows V_1 and V_2 show the direction of propagation of the two waves. Vectors \mathcal{E}_1 and \mathcal{E}_2 are the maximum values of the electric vectors of the waves at point A , with \mathcal{E}_2 assumed to be less than \mathcal{E}_1 . When these vectors differ in time phase by angle α , the total electric vector \mathcal{E} rotates in the plane of the paper with the end of the vector describing an ellipse. In Fig. 9-3 the angle α is taken as 90 deg.

The ellipse of the cross field can be readily calculated from the values of \mathcal{E}_1 , \mathcal{E}_2 , ϕ , and α as follows:

Resolve \mathcal{E}_1 and \mathcal{E}_2 into x - and y -components with x parallel to, say, \mathcal{E}_1 and with y at right angles to x . Then X and Y , the x - and y -components of the total vector, are

$$X = \mathcal{E}_1 \sin \omega t + \mathcal{E}_2 \cos \phi \sin (\omega t + \alpha)$$

or

$$X = (\mathcal{E}_1 + \mathcal{E}_2 \cos \phi \cos \alpha) \sin \omega t + \mathcal{E}_2 \cos \phi \sin \alpha \cos \omega t \quad (9-1)$$

and

$$\begin{aligned} Y &= \mathcal{E}_2 \cos (\pi/2 - \phi) \sin (\omega t + \alpha) \\ Y &= \mathcal{E}_2 \sin \phi \cos \alpha \sin \omega t + \mathcal{E}_2 \sin \phi \sin \alpha \cos \omega t \end{aligned} \quad (9-2)$$

Equations (9-1) and (9-2) are parametric equations of the ellipse in which ωt is a parameter and X and Y are the x - and y -coordinates of the locus described by the end of the total vector \mathcal{E} . If the parameter ωt is eliminated from \mathcal{E} in Eqs. (9-1) and (9-2), by solving these equations for $\cos \omega t$ and $\sin \omega t$, then by squaring and adding, the result is of the form

$$AX^2 + BXY + CY^2 = 1 \quad (9-3)$$

which is an equation of an ellipse with the axes at an angle to the x - and y -coordinates in the general case. The coefficients A , B , and C are functions of \mathcal{E}_1 , \mathcal{E}_2 , ϕ , and α .

The cross field is a common phenomenon encountered when some of the primary energy is reradiated from a reflecting object which then acts as a secondary source. In general, cross field is elliptical. At some points the ellipses collapse into a line so that the cross field becomes linear. There is also a special case when the ellipse is a circle and when, therefore, the current induced in an infinitesimal dipole in the plane of polarization has the same value regardless of the orientation.

A combination of elliptical cross field with elliptical polarization of the arriving wave results in what might be called *space or ellipsoidal polarization*.

The polarizations discussed cover the types commonly encountered.

9-3. Direction-finder Requirements above 50 Mc.—At frequencies below a few megacycles, r-f energy is propagated along the surface of the earth with least attenuation when the electric field is vertical. For this reason most low-frequency radio stations are designed to radiate vertically polarized waves. Thus, a low-frequency direction finder will, in general, be satisfactory if designed to operate only on vertically polarized waves. At higher frequencies the advantages of vertical polarization are less, and transmitters radiating horizontally polarized waves are more common.

At distances of a few hundred miles or more from a transmitter much or all of the received energy may come by way of ionosphere reflections. This effect is most pronounced on frequencies between 1 and 20 Mc. Under these conditions the polarization of the received energy may be plane or elliptical. Owing to multiple paths and variations in transmission conditions, the type of polarization of the received wave may vary with time and the polarization may pass through vertical often enough to permit bearings to be taken with a vertically responsive direction finder. It has been customary, therefore, to use vertically polarized direction finders up to about 25 Mc.

The direction-finder problem at frequencies above 50 Mc is different from that at lower frequencies because of the nature of the propagation of u-h-f waves, the relative constancy of polarization, and the effect of surrounding objects. As the frequency is increased, the wavelength becomes first comparable to and then shorter than objects of convenient dimensions, enabling the designer to make use of structures comparable in size with one-half or even with many wavelengths of the received frequency, with the result that a number of fundamentally different structures become practical and therefore come into the realm of choice.

Waves of frequencies above 50 Mc are reflected by the ionosphere only under rare conditions. Except for a moderate effect of refraction in the atmosphere, the propagation of u-h-f waves in the first approximation is a relatively simple phenomenon. At the receiving station two waves are received, a direct wave which travels through the atmosphere, and a reflected wave which after reflection from the ground or sea joins with the direct wave at the collector. The change of phase at the point of reflection at the surface of the earth or sea is usually about 180 deg when the impinging wave makes a small angle with the reflecting surface. The field strength at the receiving antenna is the vector sum of the direct and reflected fields.

One result of the interference between the direct and reflected waves is that the field strength at the receiving antenna varies with height, being very small at the surface of the ground, increasing at first almost linearly with height, and then oscillating back and forth between a low

and a high value as the height is further increased. This simple picture of propagation of u-h-f waves is made somewhat more complicated by the fact that the earth is not flat, but curved, and also by the fact that at very high frequencies, under certain conditions, the refraction of waves in the atmosphere tends to curve the waves to a degree such that the simple picture is no longer applicable. These atmospheric conditions, however, do not occur all the time and can be regarded as special.

The nature of the propagation of u-h-f waves is such that a wave originating at a transmitter as a horizontally polarized wave generally arrives at the receiver as a horizontally polarized wave. Similarly, a vertically polarized wave continues vertically polarized. There is no known phenomenon at ultrahigh frequencies that corresponds to the rotation of polarization in the ionosphere. For this reason, at ultrahigh frequencies it is not possible to use a vertical Adcock antenna (two vertical out-of-phase elements) as a universal device for determining the directions of waves of any polarization.

One of the first requirements of a universal direction finder for ultrahigh frequencies is that it be designed to receive both vertically and horizontally polarized radiation. Both horizontally and vertically polarized waves are widely used at ultrahigh frequencies, because in some cases horizontally polarized waves are preferable while in others vertically polarized waves are desirable. This means that a universal direction finder must be capable of receiving both. In fact, the ability of a direction finder to determine not only the direction of a source, but also its polarization may be of assistance in identifying the source.

The problem of direction-finder siting at ultrahigh frequencies is radically different from the corresponding problem at lower frequencies. At low frequencies the effect of conductivity in the immediate neighborhood of the collector is an important factor. At ultrahigh frequencies, objects capable of reflecting u-h-f waves are important, while the conductivity of the ground is usually a negligible factor. At low frequencies, substantial reradiation can come only from closed circuits or from large objects, *e.g.*, tall masts. At ultrahigh frequencies any objects, particularly metal ones, having dimensions in the direction of polarization greater than $\frac{1}{2}$ wavelength may be serious sources of reradiation, or reflection.

At low frequencies, of the order of 1 Mc, the field strength at the receiving site does not vary substantially with the height above the ground so that reradiating objects, elevated or not, have approximately the same effect. At ultrahigh frequencies a reflecting object placed near the ground, with the collector many feet above the ground, has a relatively minor effect because the field impinging on the reflecting object is relatively small in comparison with the field in which the collector itself is immersed. The reverse of this is also true. If a direction-finder

antenna is close to the ground, while the reflecting object is many feet above it in a much stronger field, the effect of the reflecting object is disproportionately great.

9-4. Basic Direction-finder Methods.—A *wave or phase front* may be defined as the surface along which the time phase of the electric vector is the same at a given instant. The phase front is spherical for a single point source in free space, and the direction of the electric vector is tangent to the surface of the sphere. Consider a horizontally polarized source in the plane of the page, and consider the page to be horizontal. At a point in this plane that is at a long distance from the source, the direction of the source can be determined by several different methods:

1. An infinitesimal horizontal dipole could be used as a receiving antenna which is turned to the position of maximum response, as shown in Fig. 9-4a. This orientation of the dipole would coincide with the direction of the electric vector, which is perpendicular to the direction of the source from the dipole.

2. Two very small horizontal loops placed close together in the plane of the page can be connected in phase and used as a receiving array. Since each small loop is nondirectional in the plane of the page, the maximum response is obtained when the currents induced in the two loops are in phase. This occurs when the line joining the centers of the two loops coincides with the phase front, as shown in Fig. 9-4b. If the source is vertically polarized, the corresponding receiving array could consist of two vertical dipoles connected in phase, as shown in Fig. 9-4b'. Methods 1 and 2 are different but in this case give identical results.

3. Instead of using two loops, an array consisting of a number of loops arranged in a straight line and connected in phase could be used, as in Fig. 9-4c. The maximum response would again be obtained when the line joining the loops coincides with the phase front. In case the source is vertically polarized, the corresponding receiving array could consist of a number of vertical inphase dipoles, as in Fig. 9-4c'.

The distinction between the three methods is that Method 1 indicates the direction of the electric vector at a point, Method 2 indicates the direction of the phase front in the immediate neighborhood of a point, and Method 3 indicates the direction of the phase front as averaged over some distance on both sides of a point. All three methods give identical results when there is one source in free space, but, in general, give different results when the wave front is disturbed by a reflecting object. Method 3 tends to average out the effect of ripples in the phase front caused by a reflecting object, while Methods 1 and 2 do not.

9-5. Reflections.—In conditions where the uniform radial flow of electromagnetic energy is interfered with by reflecting or absorbing objects, the assumption that the source is in a direction perpendicular to either the phase front or the electric field may not hold.

In order to discuss the conditions that exist when a wave front is

disturbed by a reflecting object, it will be convenient to distinguish between two cases. In the first case the conditions are such that the phase front is disturbed but the direction of the electric vector remains unaltered and no cross field is present. In the second case both the phase front and the direction of the electric vector are affected and cross field is present.

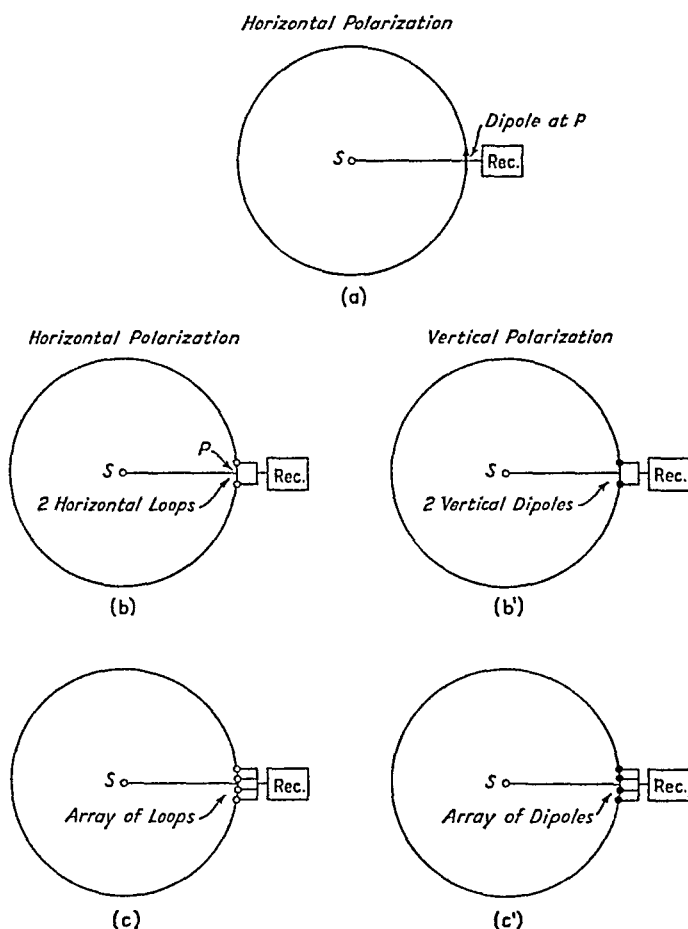


FIG. 9-4.—Basic direction-finder methods: (a) infinitesimal dipole at P ; (b) 2 horizontal inphase infinitesimal loops near P ; (b') 2 vertical inphase infinitesimal dipoles near P ; (c) several horizontal inphase infinitesimal loops; (c') several vertical inphase infinitesimal dipoles. The source is at S .

9-6. Reflections. Case 1.—Consider a vertically polarized primary source in the plane of the page and at a great distance, as indicated in Fig. 9-5, so that the field in the region of the figure is substantially uniform. The page is assumed to be horizontal and the energy to be flowing up the page. For undisturbed conditions, a phase front is as shown by the straight line from A to C . A reflecting object, such as a vertical

mast, is located at point *D*. This object absorbs some of the energy from the primary wave and is assumed to reradiate this energy as a vertically polarized wave, uniformly in all directions in the plane of the page. The mast thus acts as a secondary source, and at some distance the secondary field varies inversely with the distance from its center. It is assumed that the secondary field is equal to one-half the primary field at a distance of 4 wavelengths from the secondary source. This is the condition at point *B*. It is further assumed that the two fields are 180 deg out of phase at the secondary source. With these assumptions

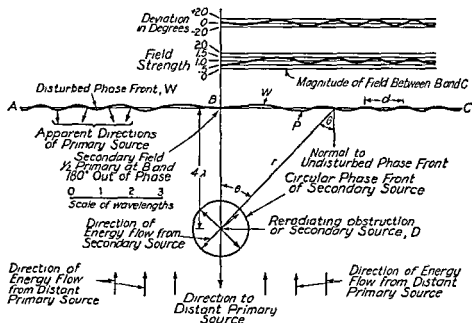


FIG. 9-5.—Illustration for Reflections, Case 1 with vertically polarized primary source at a long distance in the plane of the page and vertical mast or reflecting object at *D*.

the sum of the primary and secondary fields at any point in the vicinity of the secondary source can be expressed in the form

$$\mathcal{E} = A \sin (\omega t - \Delta)$$

where *A* and Δ are functions of the radius vector *r* from the secondary source to the point and of the angle θ between the radius vector and the normal to the primary wave front. The rippled phase front *W* in Fig. 9-5 is the locus of points for which Δ has a constant value equal to that at point *B*. As shown in Fig. 9-5, the number of ripples per unit length increases with the distance from point *B*. Near point *B* the distance between the successive crests is several wavelengths, while at large distances it approaches 1 wavelength.

If Method 2, making use of two vertical dipoles connected in phase, is

employed to indicate the direction of the phase front in the neighborhood of a point, the maximum response will be obtained when the line joining the dipoles is coincident with the tangent to the rippled phase front. In general, the direction of the normal to the rippled phase front does not pass through the distant primary source but instead may deviate by as much as 20 deg from this direction. The curve in the upper right-hand part of Fig. 9-5 shows the deviations along the phase front between *B* and *C* that would be observed by the direction-finder antenna of Method 2. The deviation indicates the difference in angular position between a normal to the disturbed phase front and to the undisturbed phase front. It is apparent that the deviations fluctuate between positive and negative values which gradually decrease with increase in distance from *B*.

If Method 3 be substituted for Method 2, using vertical dipoles distributed along a horizontal line and connected in phase, then for maximum response the array will assume a position that on the average is parallel to the phase front. If the array is several wavelengths long, the effect of the individual ripples in the phase front tends to average out and the array assumes a position nearly the same as that for the undisturbed primary phase front. If the array, for example, is about 1.5 wavelengths long horizontally and is located about 2 wavelengths to the right or left of point *B* in Fig. 9-5, a deviation will still exist since the phase front is displaced in one direction over a distance of about 2 wavelengths to one side of *B*, and a deviation of a number of degrees in the apparent direction of the primary source would result. To reduce the deviation in the vicinity of *B*, an array 3 or more wavelengths long is required. At distances of the order of 3 wavelengths or more from the point *B*, an array 2 wavelengths long will substantially reduce the effect of the ripples.

While it was convenient to plot the phase front in Fig. 9-5 in terms of wavelengths from the point *B*, this figure does not emphasize one very essential fact, *viz.*, that the number of ripples in the phase front per unit length is a function of the angle θ between the normal to the undisturbed phase front and the direction of the reflecting object. When θ is greater than about 30 deg, the approximate relation is given by

$$d = \frac{\lambda}{\sin \theta} \quad (9-4)$$

where *d* is the distance between the successive crests measured along a line parallel to *AC*. The relation holds regardless of the distance of the phase front from the reflecting object, provided that this distance is greater than several wavelengths.

No mention has yet been made of the fact that, in addition to disturbances to the phase front, the maximum value of the field also varies

from point to point. Near point B the field varies over a range of about 3 to 1. As one proceeds along the line AC away from point B the ratio of maximum to minimum field becomes less, approaching unity at a distance of many wavelengths. A curve showing the variation in the value of the field is inserted in Fig. 9-5 (second from top in upper right-hand corner).

An interesting result of the variation or gradient in field intensity occurs in the case of a direction-finder antenna consisting of two vertical closely spaced elements connected in phase opposition as in an Adecock antenna. If this antenna is placed, for example, at point P in Fig. 9-5 and adjusted to give a minimum response, it will assume a position substantially parallel to the disturbed wave front. However, the null or minimum may not be sharp, since owing to the field gradient the difference of the voltages induced in the two elements is not zero.

9-7. Reflections. Case 2.—Consider a horizontally polarized source in the plane of the page and at a great distance, as in Fig. 9-6. Consider

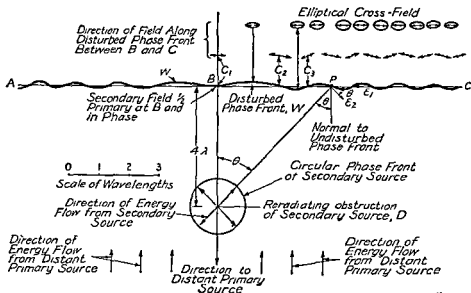


FIG. 9-6.—Illustration for Reflections, Case 2 with horizontally polarized primary source at a long distance in the plane of the page and a reflecting object at D .

also that the page is horizontal. A reflecting object located at D reradiates horizontally polarized waves uniformly in all directions in the plane of the page. We shall consider the conditions along the disturbed phase front W . At a point P the total electric field is the sum of the primary electric field ϵ_1 and the reflected field ϵ_2 . The direction of ϵ_1 is along the undisturbed phase front AC . The direction of ϵ_2 makes an angle θ with ϵ_1 .

This is a typical case of cross field (Sec. 9-2). Not only is cross field

present, but the phase front is distorted in much the same way as in Fig. 9-5. In general, the resultant electric vector describes an ellipse in the plane of the page. At points such as C_1 , C_2 , and C_3 , the ellipse of cross field collapses into a straight line because the vectors \mathcal{E}_1 and \mathcal{E}_2 are in 0- or 180-deg time phase. Elliptical cross field is present at the intermediate points.

When Method 1 is applied under the conditions of Fig. 9-6, the dipole produces maximum response when it is oriented parallel to the major axis of the ellipse of cross field. The minimum response is not zero except at points such as C_1 , C_2 , and C_3 . Method 3, using an array of horizontal loops, tends to average out the ripples, approaching closer and closer to the direction of the undisturbed wave front as the aperture of the array is increased. When an array of collinear horizontal inphase dipoles is used according to Method 3, the averaging phenomenon is more complex because not only the phase but also the direction of the total electric field at different points along the array is involved in the averaging. The result of this averaging, however, is still that the line of the array approaches the direction of the undisturbed phase front as the linear aperture or length of the array is increased.

9-8. Directivity.—Collectors, such as a half-wave antenna, an array of half-wave antennas, an electromagnetic horn, or an antenna with a concentrating reflector, may be considered to be equivalent to an array of infinitesimal dipoles or loops. If the effective aperture of such collectors is great enough, the effect of the ripples in the phase front tends to average out. There is an alternate way to understand this phenomenon. The aperture of an array is intimately related to the directivity of the array, the beam width being inversely proportional to the linear aperture. As the aperture is increased, the radiation characteristic of the array becomes more and more directional so that the effect of secondary or reflected radiation coming from a direction to the side of the main lobe is reduced as the array aperture is increased.

If the aperture of the collector is sufficiently great, its directivity or resolution may be sharp enough to allow it to distinguish between the energy arriving directly and via the reflecting or reradiating object. Thus, directivity may be a distinct advantage in reducing bearing deviations due to reflecting objects.

9-9. Vertically Polarized Collector in Horizontally Polarized Primary Field.—The effects of a reflecting object are not limited to the types of disturbance that have been discussed above. There is another phenomenon that may be encountered when reflecting objects disturb the primary field. Consider a distant horizontally polarized source S in the plane of the page, as in Fig. 9-7. The page is assumed to be horizontal, and the waves from the source are assumed to be traveling up the page. Reflect-

ing object R is inclined to the plane of polarization so that the reradiated field has vertically as well as horizontally polarized components. It will be assumed that a vertical Adcock antenna is used to determine the direction of the primary source. Since the primary source is horizontally polarized, substantially no energy is received from it directly by the Adcock antenna. The energy that is received is that converted into vertically polarized waves by the inclined reflecting object. In effect there is only one source that is vertically polarized, *viz.*, the reflecting object, and the bearing obtained is that of this object and not of the primary source. If a horizontally polarized collector were used, the vertically polarized component from the reflecting object would be rejected and the

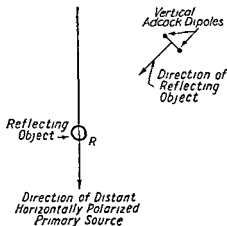
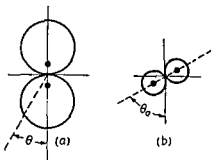


FIG. 9-7.—Vertically polarized direction-finding antenna in horizontally polarized primary field and secondary field with vertically and horizontally polarized components.

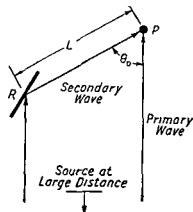


FIG. 9-8.—Antenna at point P in primary and secondary fields arriving from directions making an angle θ_0 . Fig 9-9 shows the result of the two fields on the pattern of a small vertical loop.

phenomenon encountered would be the same as that met with in connection with Reflections, Case 2.

9-10. Effect of Reflections on Antenna Pattern.—The effect of a reflecting object on the phase front has been considered. It may also be instructive to treat the problem from another point of view, *viz.*, from the effect on the antenna pattern.

Consider a vertically polarized primary source in the plane of the page at a great distance, as in Fig. 9-8. The page is assumed to be horizontal. Consider a small vertical loop-direction-finder antenna at the point P , with a directional characteristic in the horizontal plane given

by $\cos \theta$ where θ is the direction of arrival of energy with respect to the plane of the loop, as in Fig. 9-8a. Assume that a reflecting object at R reflects some of the direct radiation from the primary source toward P , thus becoming a secondary source of radiation, of the same polarization. The angle between the direction of arrival of the energy from the primary

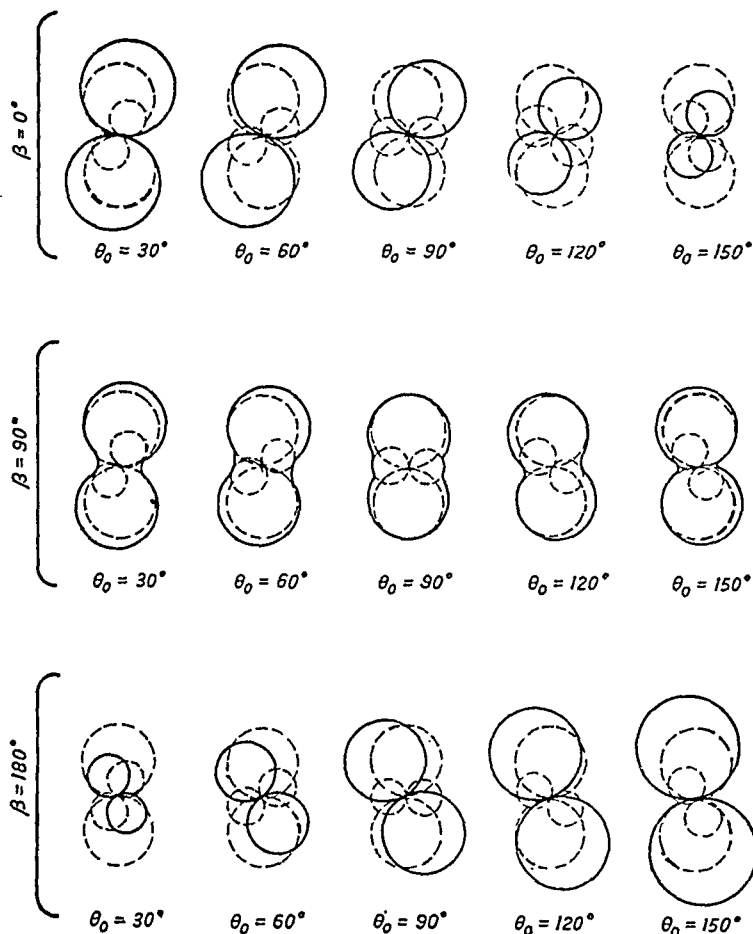


FIG. 9-9.—Calculated loop-antenna patterns showing the effect of a secondary field produced by a reflecting object as shown in Fig. 9-8. The secondary field arrives from a direction making an angle of θ_0 with respect to the direction of the primary field. The two fields have a time-phase difference β and the secondary field is one-half the primary field at the loop antenna. The large dashed curves are the patterns for the primary field, the small dashed curves the patterns for the secondary field, and the solid curves the resultant patterns.

and secondary sources at P is θ_0 . It is assumed that the distance L between R and P is sufficiently large so that the secondary wave front can be considered plane over the extent of the loop antenna. Consider also that the magnitude of the secondary field is one-half that of the primary at point P . If only the secondary field were present, the resulting

loop-antenna directional pattern would be smaller and displaced by an angle θ_0 , as shown in Fig. 9-8b.

With both primary and secondary fields present at P , the resulting loop-antenna pattern is a function of both the angle θ_0 and the time phase β between the primary and secondary fields. The resulting patterns for values of the angle θ_0 of 30, 60, 90, 120, and 150 deg and values of β of 0, 90, and 180 deg are presented in Fig. 9-9. The pattern calculations are made by adding the response to the two sources with proper attention to magnitude and phase. The large dashed curves are the patterns for the primary field only and the small dashed curves for the secondary field only. The solid curves show the patterns resulting from both fields.

It is apparent that in most cases the resultant loop pattern has an angular displacement, indicating a disturbed phase front. The maximum displacement is about 30 deg for the case of $\theta_0 = 120^\circ$ and $\beta = 0^\circ$ (or $\theta_0 = 60^\circ$ and $\beta = 180^\circ$).

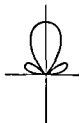


FIG. 9-10 —
Three-lobed antenna pattern.

When β is 0 or 180 deg, the shape of the pattern is unchanged so that both the maximums and minimums are displaced by the same angle. Thus in using the loop as a direction-finding antenna the same deviation of the bearing of the primary source would be observed, whether the loop were adjusted for maximum or minimum response. However, it should be noted that this is the case only when the time phase β of the primary and secondary fields is 0 or 180 deg. When β is 90 deg, the pattern shape is changed and the minimums are no longer deep and sharp.

In this discussion, the patterns of Fig. 9-9 are considered to be those of a small vertical loop in a vertically polarized field. The same patterns also apply to the case of an infinitesimal horizontal dipole in a horizontally polarized field.

It is interesting to note the effect of a secondary field on the pattern of another type of antenna. For example, let us consider an antenna with a directional characteristic that has a main lobe and two small side lobes, as shown in Fig. 9-10. (The phase of the side-lobe response is assumed to be 180 deg with respect to the main lobe.) Consider that an antenna with such a three-lobed pattern is placed at the point P of Fig. 9-8 and that the secondary field is one-half the primary. The patterns resulting from the combined effect of the primary and secondary fields are presented in Fig. 9-11 for values of θ_0 equal to 30, 60, and 90 deg and of β equal to 0, 90, and 180 deg. In general the effect of the secondary field on the resultant pattern is to produce both a change in shape and an angular displacement. It is to be noted that the angular displacement of the main lobe is greatest when θ is 30 deg. When θ_0 is 60 or 90 deg,

the angular displacement of the main lobe is relatively small although the shape may be distorted.

9-11. Time-difference Method.—In the preceding discussion, direction-finding methods involving the direction of the electric field and of

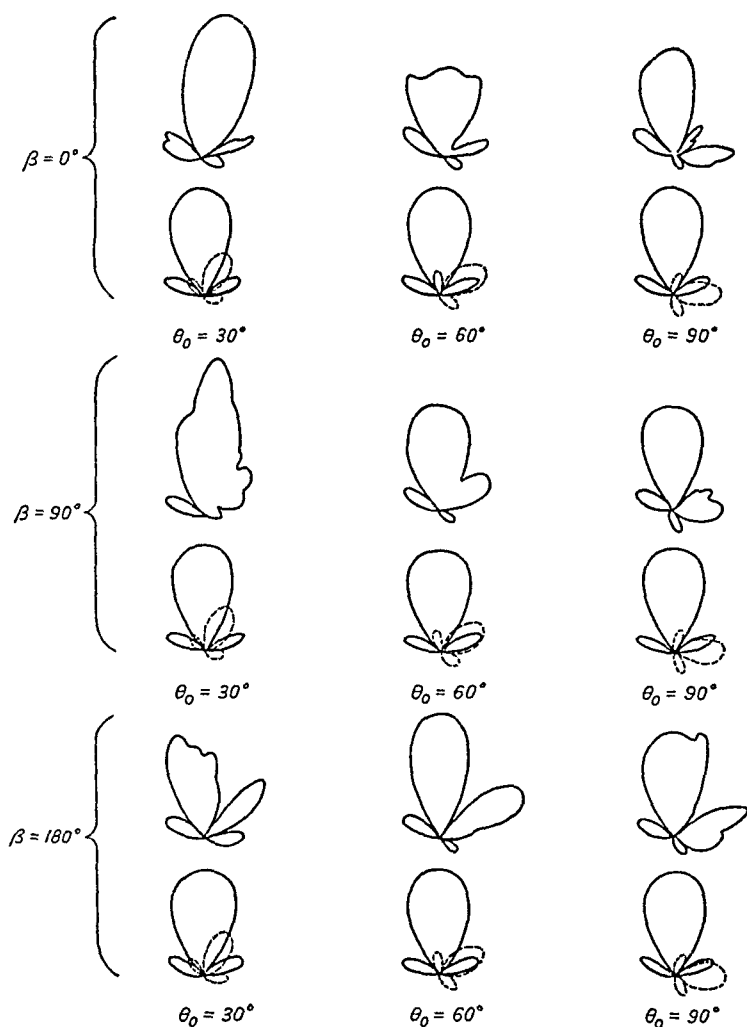


FIG. 9-11.—Calculated patterns for antenna with normal 3-lobed characteristic showing the effect of a secondary field produced by a reflecting object, the direction of which makes an angle θ_0 with respect to the direction of the primary field. The two fields have a time-phase difference β and the secondary field is one-half the primary at the antenna. The upper curve of each set indicates the pattern for the combined fields. The lower curves show the pattern for the primary field alone (solid) and for the secondary field alone (dashed).

the phase front have been treated. A basically different method of direction finding makes use of the time-difference principle. This method is chiefly of interest when the radio energy is transmitted in the form of short pulses. An arrangement employing this method is illus-

trated in Fig. 9-12. A distant source in the direction shown radiates a pulse-modulated carrier. A_1 and A_2 are two nondirectional antennas, R_1 and R_2 are two receivers, V_1 is a video inverter, P_1 is a calibrated variable-phase delay, and P_2 a fixed-phase delay. The output of receiver

R_1 is inverted and added to the output of receiver R_2 . The sum of the outputs is observed on a synchronized cathode-ray oscilloscope. When pulses received by R_1 and R_2 arrive simultaneously and the time delays P_1 and P_2 are adjusted to equality, no signal is observed. When the variable phase delay P_1 is decreased, a picture similar to that in Fig. 9-12b is observed. When time delay P_1 is increased, a picture similar to that in Fig. 9-12c is observed. When the distant source is located so that the phase front makes an angle ϕ with the line joining antennas A_1 and A_2 , antenna A_1 receives the pulse later than antenna A_2 . A balanced condition is present halfway between the conditions shown in Figs. 9-12b and 9-12c and occurs when the time delay P_1 is decreased by the required amount. Time delays in the receivers are calibrated out by the use of a local pulsed source placed halfway between antennas A_1 and

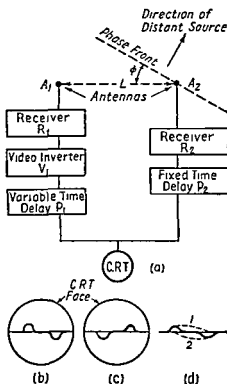


FIG. 9-12.—(a) Arrangement for direction-finding by time-difference method; (b) and (c) presentation on c-r-t face; (d) addition of pulses 1 and 2 to give sum shown at (b).

A_2 . The time difference between the arrival of the phase front at antenna A_2 and at antenna A_1 is given by

$$t = \frac{L \sin \phi}{300} \quad (9-5)$$

where t = the time difference, microseconds

L = the distance between the antennas, meters

ϕ = the angle between the phase front and the line between the antennas, as shown in Fig. 9-12.

From Eq. (9-5) it is apparent that the time difference and hence the bearing accuracy is proportional to the distance L between the antennas.

Since the change in $\sin \phi$ with respect to ϕ is greatest when ϕ is zero

$$\left[\frac{d(\sin \phi)}{d\phi} = \cos \phi \right]$$

the bearing accuracy is greatest when the source is in a direction perpendicular to the line joining the antennas.

9-12. Other Methods.—At extremely high frequencies it is sometimes convenient to use a thermosensitive element at the focal point of a parabolic reflector. This method indicates the direction of the average phase front, the averaging being done by the parabolic reflector.

An electromagnetic wave has still another property that could be used to determine the direction of propagation. This property is radiation pressure. The radiation pressure, however, is very small even when the wave carries considerable energy. For this reason, practical application of this principle to radio direction finding seems remote at this time.

9-13. Indicators.—A complete direction-finder system consists of an antenna or collector, a receiver, and some means for presenting the bearing information either visually or aurally, usually called an *indicator*.

One of the simplest direction finders consists of a manually rotatable directional collector, a receiver, and a pair of headphones. Manual rotation of the collector may be satisfactory when the intensity of the source is relatively constant. If the intensity of the source changes rapidly and by a large factor, other means may be required. For

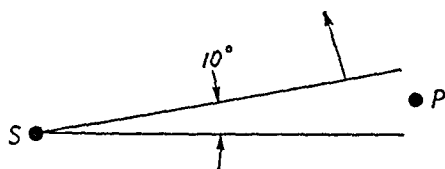


FIG. 9-13.—Source with rotating 10-deg beam.

example, consider a directional source S with a beam width of about 10 deg which rotates continuously in azimuth ten times per minute (Fig. 9-13). A sweeping source of this type is widely used for radars. Since the angular velocity is 60 deg per sec, the beam of the source sweeps by a distant point P in $\frac{1}{6}$ sec. This time interval is too short for manual measurement, although by observing successive sweeps of the source it is still possible to obtain the bearing, given sufficient time. Even such a procedure usually fails in practice since the intensity of the signal may change during consecutive sweeps, because of the motion of the source or of the direction-finder craft or vehicle. The combination of a stationary source and stationary direction finder is not common.

Because of these fundamental limitations of the manual direction finder and in order to provide means for taking bearings rapidly and sometimes simultaneously on several sources on the same frequency, a variety of semiautomatic and automatic indicators has come into use. One type of automatic indicator system consists of a rotating directional

antenna with the deflecting coils of a cathode-ray oscilloscope rotating in synchronism. The general arrangement is illustrated in Fig. 9-14. The received signal is rectified and applied to the deflecting coils of a cathode-ray tube. As the directional receiving antenna sweeps through the direction of the source *S* the spot on the cathode-ray-tube face traces out a curve similar to the directional pattern of the antenna, with the maximum deflection usually indicating the direction of the source, relative to some reference direction. In the example of Fig. 9-14 the source is at a bearing of 90 deg, or to the east.

The information obtained by the direction-finder collector may, if

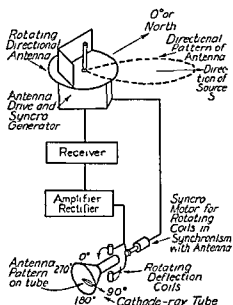


FIG. 9-14.—Basic arrangement of rotating-antenna direction-finder system.

desired, be resolved into two components, (1) the r-f power received, and (2) the angular position or bearing from which it is received. In the system shown in Fig. 9-14 the radial displacement of the cathode-ray beam is some function of the power received, and the angular displacement indicates the direction in which the antenna is pointed. The radial and angular displacements combine on the oscilloscope to yield a pattern from which the bearing of the source can be deduced.

Since with this type of presentation it is not necessarily the point of maximum deflection of the pattern that indicates the direction it is often the practice to bisect the pattern to obtain the direction of the source. In this way the entire pattern including maximum and sides is used. Even though the directional pattern is not extremely sharp, the pattern can be bisected with considerable accuracy provided the pattern is symmetrical.

A widely used modification of the system of Fig. 9-14 incorporates an electrostatic-deflection type of cathode-ray tube instead of the magnetic type. Both types were developed at Radio Research Laboratory. In the electrostatic type the amplified receiver output is applied to the rotor of a so-called *scanning capacitor*. The rotor of the capacitor turns in synchronism with the antenna. As the rotor turns, two voltages in quadrature are induced on two sets of stator plates. These voltages are then applied to the horizontal and vertical deflecting plates of an electrostatic cathode-ray tube.

With the electrostatic-deflection system the individual pulses of a pulse-modulated source can be presented (see Chap. 11). Since there usually are relatively long silent intervals between the pulses of a single source, the pulses from other sources can be presented during these intervals. Thus, two pulse-modulated sources on the same frequency and in the same or different directions can be presented simultaneously and the bearings of both sources obtained.

A number of antennas for use with these systems and also other types are described in Chap. 10, together with a discussion of their principles of operation. Indicators for a variety of systems are described in Chap. 11, and a more detailed discussion of the above types is included.

CHAPTER 10

ANTENNAS FOR DIRECTION FINDERS

By J. D. KRAUS, H. K. CLARK, E. C. BARKOFSKY, AND G. STAVIS

10-1. General Considerations.—A radio antenna may be described as a device for converting guided energy into radiant energy, or vice versa. An antenna acts as a transformer between space on the one hand, and a transmission line or waveguide on the other. A direction-finder antenna is a receiving antenna or collector, and its function is both to extract energy from passing electromagnetic waves and to indicate the direction from which the energy arrives. In this chapter antenna arrangements for direction finders will be discussed with particular reference to those suited for frequencies of 50 Mc and higher. Receivers and

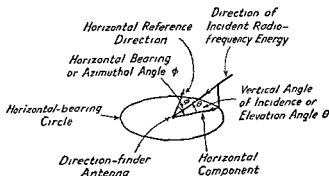


FIG 10-1 —Relation of direction angles.

indicators suited for use with these antennas to form complete direction-finder systems are discussed in other chapters.

In the most general case, a direction finder should be able to indicate the direction of arrival of energy originating from any point in three-dimensional space. In the following discussion the problem is simplified by requiring the direction finder to indicate only the directional component in a horizontal plane.

The terms used to describe the geometrical relations between the direction of arrival of the incident energy and the reference planes are illustrated in Fig. 10-1. Assume a direction-finder antenna located at the center of the horizontal bearing circle as shown. The bearing angle ϕ is normally read in a clockwise direction from an arbitrary reference

direction taken as "0" deg. For *true bearings*, this reference direction is true north.

The vertical or elevation angle θ is taken as positive above the plane of the circle and as negative below. For direction finders located near the surface of the earth, the elevation angles will, in general, be positive. If the source is distant, this angle will be small. For airborne direction finders the elevation angle will, in general, be negative when the sources are on the surface of the earth, but may be either positive or negative for sources located in other aircraft.

Although the elevation angle of the incident energy may not be zero, the problem is simplified in the direction finders to be described, by requiring the equipment to indicate only the bearing angle.

10-2. Beam Width.—With the exception of antennas for direction finders using the time-difference method described in the preceding chapter, collectors or antennas for direction finders can, in general, be classified on the basis of their *directivity*. For a receiving antenna this is the property by which the antenna collects more energy from certain directions than others. A graph showing the variation in the energy received as a function of angular

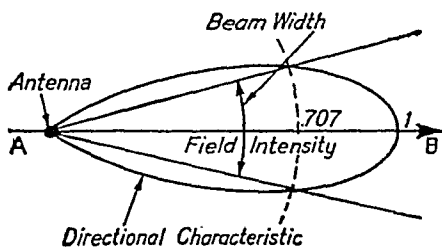


FIG. 10-2.—Half-power beam width.

direction is called a *directional characteristic* or *radiation pattern*. The directional characteristic of an antenna is the same for transmission or reception.

The directivity of an antenna may be expressed in various ways. One method is in terms of the beam width. It is customary to measure the width of the beam between half-power points, *i.e.*, the points on the directional characteristic at which the power flow per unit area is one-half the maximum value for the antenna. The beam width between these points is called the *half-power beam width* and is expressed in degrees or radians. In Fig. 10-2 the directional characteristic of an antenna is shown in terms of field intensity, with the maximum taken as unity. Since the field intensity varies as the square root of the power per unit area, the half-power beam width is between points on the radiation characteristic having $1/\sqrt{2}$ or 0.707 the maximum value.

The beam width shown in Fig. 10-2 is that in the plane of the page. The directional characteristic is actually a three-dimensional figure, and a more complete specification of beam width involves two widths, one in the plane of the page, as shown, and one in the plane perpendicular to the page, through the line AB. In the following discussion the directivity of an antenna will sometimes be designated by the beam width in two planes.

In discussing directivity it is instructive to consider the relation between the beam width, the linear aperture, and the physical aperture. As defined in Sec. 10-7 the *physical aperture* is the physical cross-sectional area of the antenna perpendicular to the direction of propagation of the incident wave, with the antenna oriented for maximum response. We may speak of the *linear aperture* as the largest linear dimension of the antenna. To illustrate the relation between beam width and aperture a number of examples of practical antennas will be cited, extending from types that are a small fraction of a wavelength in linear aperture to types many square wavelengths in physical aperture.

10-3. Antennas of Dimensions Small in Comparison with a Wavelength.—At low frequencies the electric field of interest is normally vertically polarized so that a direction finder for these frequencies should show a directional characteristic when rotated horizontally in a ver-

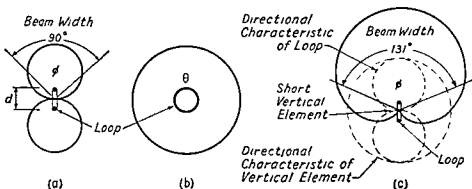


FIG. 10-3.—Directional characteristics of single vertical loop alone (a) and (b), and in combination with a vertical element (c).

tically polarized field. A convenient form of collector meeting this requirement is a single vertical loop. The directional characteristic of a single-loop antenna, the diameter of which is a small fraction of a wavelength, is shown in Fig. 10-3a. The pattern is bidirectional and in the shape of a figure of eight. The half-power beam width in the horizontal plane (a function of ϕ) is 90° . Rotation of the loop in its own plane produces no directional effects so that the beam width in this plane may be called 360° , as shown in Fig. 10-3b. When the loop diameter is a small fraction of a wavelength, the radiation resistance is small and the sensitivity may be low even though a number of turns are used. For this reason the loop is sometimes tuned to resonance. In the region of the minimum or null in the horizontal-plane pattern, the response of the loop varies rapidly with angle ϕ . The null region is, therefore, commonly used for directional indications. If the loop diameter is of the same order as the wavelength, the pattern is not like that of Fig. 10-3a but has several lobes, which may produce confusion when the null method is used.

Another directional arrangement frequently used at low frequencies is the combination of a single vertical loop and a short vertical element. The loop pattern is a figure of eight, and the vertical element pattern is a circle. When the loop and vertical element voltages are equal and added in phase, a cardioid response results as shown in Fig. 10-3c.

Other more directional antennas, the dimensions of which are a small fraction of a wavelength, can be constructed. For example, two vertical out-of-phase loops placed side by side with their planes parallel yield a four-lobed pattern with a beam width of about 45 deg, as shown in Fig. 10-4a. Four vertical infinitesimal dipoles are shown in Fig. 10-4b at the corners of a square, with the dipoles at the corners connected by one diagonal in opposite phase to those connected by the other diagonal. The pattern has four lobes. Four dipoles arranged in this way form what may be called an *octapole*. Combinations of larger numbers of dipoles can produce patterns of still more lobes of yet smaller beam width. Although these and other highly directional structures of small aperture may be constructed, their practical application is limited by their extremely low radiation resistance. For this reason it is convenient arbitrarily to restrict our discussion of antennas of small aperture to one of the most practical types, *viz*, the single loop.

10-4. Antennas about $\frac{1}{2}$ Wavelength Long.—If structures about $\frac{1}{2}$ wavelength long can be constructed, as is often the case at the higher frequencies, the designer has a choice of several types of practical antennas. Four simple varieties will be described in this class or group.

For horizontal polarization the simplest type is a linear half-wave element. The horizontal beam width of this antenna, as shown in Fig. 10-5a, is about 78 deg, or 12 deg less than for a loop. When rotated about its axis, the half-wave element has no directivity, as shown in Fig. 10-5b. The beam width and the patterns of an infinitesimal horizontal dipole for horizontal polarization are the same as those of a small vertical loop for vertical polarization (see Figs. 3a and b). Horizontal end-loaded linear elements shorter than $\frac{1}{2}$ wavelength are sometimes used, with horizontal-plane beam widths having values between those of the infinitesimal dipole and the half-wave element.

Another antenna in this group is the Adcock. This consists of two linear spaced elements connected in phase opposition. The elements

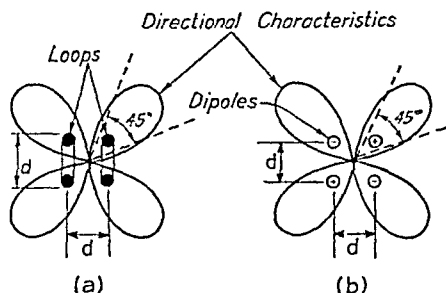


FIG. 10-4.—Directional characteristics of (a) two out-of-phase vertical loops and (b) four vertical dipoles or octapole for spacing d small compared to a wavelength.

are normally about $\frac{1}{2}$ wavelength long, and the spacing is about $\frac{1}{2}$ wavelength or less. The directional pattern of a horizontal Adcock for horizontal polarization is shown in Fig. 10-5c. The beam width in the horizontal plane or plane of the elements is about 72 deg. The pattern of a vertical Adcock for vertical polarization is illustrated in Fig. 10-5d. The beam width in the horizontal plane or plane perpendicular to the elements is about 120 deg. This antenna has directivity in two planes,

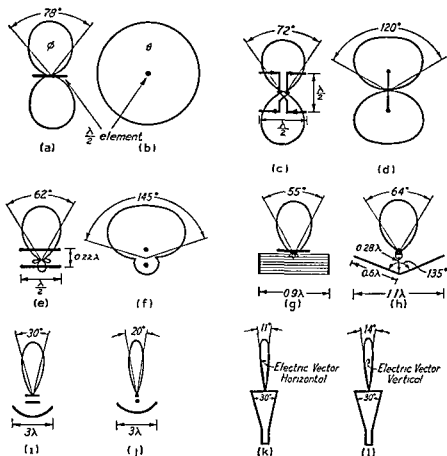


FIG. 10-5.—Decrease in beam width with increasing aperture.

whereas the simple loop and dipole possess directivity in only one. For this reason the Adcock may be used vertically for direction finding in vertically polarized fields or horizontally in horizontally polarized fields.

A third antenna or collector belonging to this group consists of a half-wave element with a half-wave reflector. Various spacings between the element and reflector may be used. Values usually lie between 0.1 and 0.4 wavelength. An example of an array of this type with 0.22-wavelength spacing is illustrated in Figs. 10-5e and f. The half-wave

element in this example is of the sleeve type, described in Chap. 5. With the elements horizontal the beam width for horizontal polarization is about 62 deg (Fig. 10-5e), and with the elements vertical the beam width for vertical polarization is about 145 deg (Fig. 10-5f). With the exception of the cardioid arrangement, this is the first antenna we have discussed which has a unidirectional pattern, *i.e.*, one in which the radiation characteristic in one direction is greater than in any other.

A fourth example in this group is an antenna with corner reflector. This antenna is commonly constructed with maximum dimensions between $\frac{1}{2}$ and 1 or 2 wavelengths. The directional characteristics of a half-wave sleeve element with a corner reflector of 135-deg angle and an aperture of 0.9 by 1.1 wavelength are illustrated in Figs. 10-5g and h. The beam widths for this antenna are 55 and 64 deg. This antenna also has a unidirectional characteristic and the narrowest beam width of any discussed thus far. The antenna with corner reflector can be applied to either vertically or horizontally polarized waves by turning the entire array so that the linear element is parallel to the electric field in question.

10-5. Antennas with Dimensions of Several Wavelengths.—If linear aperture of several wavelengths or more can be constructed, still other types of antennas become candidates. Examples are electromagnetic horns and antennas with parabolic reflectors. With these antennas the beam width is to a first approximation inversely proportional to the linear aperture in the plane in which the beam width is measured. An example of a primary antenna with parabolic reflector is presented in Figs. 10-5i and j. The reflector in this case is a parabola of revolution or paraboloid with an aperture 3 wavelengths across. The beam widths for this antenna are about 20 and 30 deg. A horn antenna with a square aperture about 5 wavelengths on a side and 30-deg flare is illustrated in Figs. 10-5k and l. The beam widths for this horn are about 11 and 14 deg. An antenna consisting of an array of a large number of half-wave elements might be included in this group. It is omitted here, however, as being a relatively narrow-band device, since in practice the proper phasing of the elements can be maintained only over a relatively narrow band. The antennas described above are better suited for relatively wide-band applications.

With antennas having a unidirectional pattern, such as a corner reflector, parabolic reflector, and horn, either the maximum of the pattern or the entire pattern is used for the directional indication. In general the nulls or minimums of these patterns have no fixed relation to the antenna axis over any appreciable frequency range.

10-6. Beam Width vs. Aperture.—The foregoing antennas have been considered as belonging to one of three groups, which may now be redefined as follows:

1. Antennas with linear apertures that are a small fraction of a wavelength
2. Antennas with linear apertures of the order of 1 wavelength
3. Antennas with linear apertures of many wavelengths

For most practical purposes, directional antennas in Group 1 are of the loop type. In Group 2 there is a greater variety ranging from the half-wave element to an antenna with a corner reflector. Group 3 includes horns and parabolic reflector antennas.

It is obvious that antennas in Groups 1 and 2 could be used wherever structures of larger apertures are applicable, but it is assumed that for

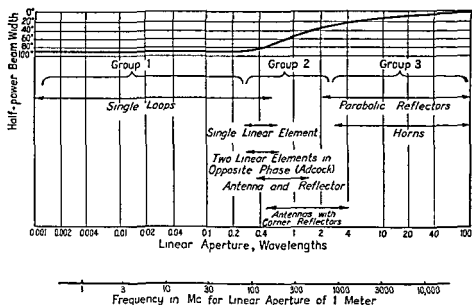


FIG. 10-6 — Beam width of representative types of antennas as a function of linear aperture in wavelengths.

most direction-finder applications the narrowest possible beam is desired. An exception to this is that a wide beam in the vertical plane may sometimes be desired in order to accept energy arriving at considerable vertical angles of incidence. Such a beam may also be desirable to allow for rolling or rocking of the craft or vehicle carrying the direction finder. Thus a fan-shaped beam as narrow as possible in the horizontal or ϕ direction (see Fig. 10-1) and 20 to 50 deg wide in the vertical or θ direction may be called for. To fulfill this requirement, antenna structures that are wide horizontally and relatively narrow vertically are commonly employed.

In the examples shown in Figs. 10-3 and 10-5, there is a trend toward narrower beam widths in both planes as the physical aperture area in square wavelengths is increased. There is also a trend toward a more unidirectional characteristic as contrasted to a bidirectional character-

istic. The over-all trend is, thus, toward a concentration of the energy more and more in a single direction if the antennas are transmitting energy. When the antennas are used as collectors, the trend is toward increasing the amount of energy received from one direction relative to that received from all others.

In Fig. 10-6, antennas of the three groups discussed above are presented as a function of the linear aperture in wavelengths. The frequency limits over which a given antenna is considered suitable are arbitrary. At the top of the figure a graph is inserted showing the trend in approximate minimum beam width of the various antennas as a function of the linear aperture in wavelengths. A single loop (Group 1) has a beam width of 90 deg when the loop is a small fraction of a wavelength across. In general, antennas in Group 2 have narrower beams. Antennas in Group 3 may have very narrow beams. For example, with a 10-wavelength linear aperture a beam width of less than 10 deg is obtainable, while a beam width of less than 1 deg may be had for a linear aperture of 100 wavelengths.

To summarize: For linear apertures of less than about 0.2 wavelength, the beam width is 90 deg; for linear apertures of the order of 1 wavelength, narrower beams may be produced; and for linear apertures of several wavelengths or more, a region is entered where to a first approximation the beam width is inversely proportional to the linear aperture.

It is obvious that the particular type of antenna that is used for a given frequency depends largely on the physical size of the structure that it is practicable to build. For purposes of illustration let us arbitrarily assume that the antenna aperture is 1 m across. Then the frequency corresponding to a linear aperture of 1 wavelength is 300 Mc, as indicated by the scale inserted at the bottom of Fig. 10-6. It is apparent that for a linear aperture of 1 m a loop is a practical choice below about 50 Mc, while horns and parabolas enter the picture at frequencies above about 1000 Mc. Between 50 and 1000 Mc the various antennas of Group 2 are candidates. A linear aperture of 1 m is quite appropriate in the present discussion, since most of the direction-finding antennas described below are of about this size or smaller. An antenna of 1-m linear aperture is convenient for shipborne applications, but may be too large for airborne applications.

In the following sections, direction-finder antennas are described that were developed at the Radio Research Laboratory for use on frequencies between 50 and 12,000 Mc. The antennas include representatives of each of the three groups outlined above.

10-7. Aperture.—One requirement of a direction-finder antenna is that it possess a suitable directional characteristic. Another requirement is that it collect and deliver to the receiver sufficient energy to

permit reception of the source, the direction of which is desired. This collecting ability is sometimes referred to as the *sensitivity*. However, in this section the collecting ability will be discussed as a function of the aperture of the antenna. To facilitate our discussion of this relation it will be convenient to define three types of aperture, viz., *effective aperture*, *collecting aperture*, and *physical aperture*. In addition the *power gain* and *directivity* will be defined and the relations between these terms discussed.

10-8. Effective Aperture.—When a receiving antenna is immersed in an electromagnetic field, an effective voltage E is induced in the antenna. This voltage is a function of the power W in watts per square meter of the incident energy, its direction of arrival, the size and design of the antenna, etc. Figure 10-7a shows an antenna in a field of W watts per sq m and terminated by a very short line in an impedance Z_1 . The equivalent circuit of the antenna and terminating impedance is indicated in Fig. 10-7b, in which Z_0 is the antenna impedance as measured between the antenna terminals. The effective current I through the terminating impedance is

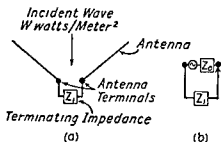


FIG. 10-7.—(a) Receiving antenna or collector with terminating impedance Z_1 ; (b) equivalent circuit.

pedance is indicated in Fig. 10-7b, in which Z_0 is the antenna impedance as measured between the antenna terminals. The effective current I through the terminating impedance is

$$I = \frac{E}{|Z_0 + Z_1|} \quad (10-1)$$

Both antenna and terminating impedances have resistive and reactive components R_0 , X_0 and R_1 , X_1 , respectively. Antenna resistance R_0 is actually the sum of the radiation resistance R_r and the loss resistance R_l .

The effective or useful power W_1 delivered to the terminating impedance is that which appears in the terminating resistance R_1

$$W_1 = I^2 R_1 \quad (10-2)$$

where

$$I = \frac{E}{\sqrt{(R_0 + R_1)^2 + (X_0 + X_1)^2}} \quad (10-3)$$

Substituting in Eq. (10-2) the values of I and of R_0 , we obtain

$$W_1 = \frac{E^2 R_1}{(R_r + R_l + R_1)^2 + (X_0 + X_1)^2} \quad (10-4)$$

The ratio between the power in the terminating impedance W_1 and the power per square meter of the incident wave W will be defined as the

effective aperture A_e of the antenna in square meters, i.e.,

$$\text{Effective aperture} = A_e = \frac{W_1}{W} \quad (10-5)$$

If W_1 is expressed in watts and W in watts per square meter, the effective aperture is in terms of square meters. If W is in terms of watts per square wavelength, however, then the effective aperture is in terms of square wavelengths.

Introducing the value of W_1 given in Eq. (10-4) we obtain for the effective aperture

$$A_e = \frac{E^2 R_1}{W[(R_r + R_e + R_1)^2 + (X_0 + X_1)^2]} \quad (10-6)$$

It is apparent from Eq. (10-6) that the effective aperture takes into account antenna losses and any mismatch between the antenna and terminating impedance. It is also apparent that the effective aperture is a maximum when the collector is oriented for maximum response.

10-9. Collecting Aperture.—Let us next consider the case in which the maximum power is transferred by the antenna to the terminating resistance. Assume that the antenna is resonant and therefore that the antenna reactance X_0 is zero. For maximum power transfer the terminating impedance must be chosen to match the antenna resistance. The terminating reactance X_1 should thus be zero and the terminating resistance equal to the antenna resistance. Let us assume further that the antenna loss resistance is zero so that the antenna resistance equals the radiation resistance R_r . With the conditions stated the maximum power in the terminating resistance becomes

$$W_1' = \frac{E^2}{4R_r} \quad (10-7)$$

This is the power delivered to the terminating impedance when the antenna has no losses and the terminating impedance has the optimum value. The ratio of this power to the power per unit area in the incident wave can be defined as the *collecting aperture* A_e . Therefore

$$A_e = \frac{W_1'}{W} = \frac{E^2}{4WR_r} \quad (10-8)$$

The collecting aperture will be in the same units as the area used in expressing W .

It is apparent that the collecting aperture is a maximum when the antenna is oriented for maximum response.

10-10. Physical Aperture.—The third type of aperture, the *physical aperture*, can be defined as the physical cross-sectional area of the antenna

perpendicular to the direction of propagation of the incident wave, with the antenna oriented for maximum response.

10-11. Relation between Apertures.—Let us now examine further the relation between the three types of apertures that have been defined. It is apparent that the collecting aperture of an antenna is the maximum value that the effective aperture can attain. Thus, the effective aperture will always be less than or equal to the collecting aperture, and the ratio of the two apertures can be called the *effectivity* ϵ . Thus,

$$\epsilon = \frac{A_e}{A_c} \quad (10-9)$$

The effectivity may range in value between 0 and 1.

The ratio of the collecting aperture to the physical aperture can be called the *collectivity* κ , thus

$$\kappa = \frac{A_c}{A_p} \quad (10-10)$$

From Eqs. (10-9) and (10-10) we obtain

$$A_e = \epsilon A_c = \epsilon \kappa A_p \quad (10-11)$$

The collecting aperture, in the case of antennas with parabolic reflectors or electromagnetic horns, is usually of the same order as the physical aperture, or somewhat less. The collecting aperture of a resonant half-wave element may, however, greatly exceed the physical aperture. Thus the collectivity κ may assume values both greater and less than 1.

10-12. Power Gain.—The collecting ability of an antenna can be expressed in terms of the *power gain* over some type of reference antenna. For example, the reference antenna may be a resonant half-wave element in free space, terminated in its optimum impedance and oriented for maximum response. The antenna under test, however, may or may not be terminated for maximum power transfer. Thus, the effect of any mismatch at the antenna termination will appear in the measured power gain. On this basis the *power gain* G can be defined as the ratio of the power delivered to the terminating impedance of the antenna under test to the power delivered to the optimum terminating impedance of a resonant half-wave element when both are immersed in a field of the same power per unit area. If the antenna under test is oriented for maximum response, the ratio gives the maximum power gain.

The effective aperture of the antenna can be directly related to the power gain provided the collecting aperture of the reference antenna is known. The collecting and effective apertures of the reference antenna are of course equal. The collecting aperture A_k of a resonant half-wave element, with a radius small compared with its length, is about 0.13

square wavelength. The effective aperture of an antenna whose power gain over a resonant half-wave element is known can thus be obtained from the relation $G = A_e/A_h$. Then

$$A_e = A_h G = 0.13G \quad (10-12)$$

in square wavelengths.

For any specific frequency the collecting aperture of the resonant half-wave element can be expressed in square meters. Computed values of the effective aperture of a resonant half-wave element are presented in Fig. 10-8 for frequencies from 50 to 10,000 Mc. The apertures vary from about 5 to 0.0001 sq m over this range. Consider for a moment the aperture at 100 Mc which is about 1.2 sq m. A half-wave is 1.5 m long at this frequency. If a 100-Mc half-wave antenna is made of $\frac{1}{4}$ -in. diameter tubing, the cross section or physical aperture is only about 0.01 sq m. The collecting aperture is 1.2 sq m, or 120 times as great. While a half-wave element may gather energy over an area much greater than its physical aperture, this is not true with many other types of antennas, especially those discussed under Group 3 above. For example, parabolic reflectors with apertures of many square wavelengths usually are found to possess collecting apertures about equal to or less than their physical apertures. In general any well-designed antenna with an aperture of many square wavelengths will be found to have a collecting aperture of about the same order as its physical aperture.

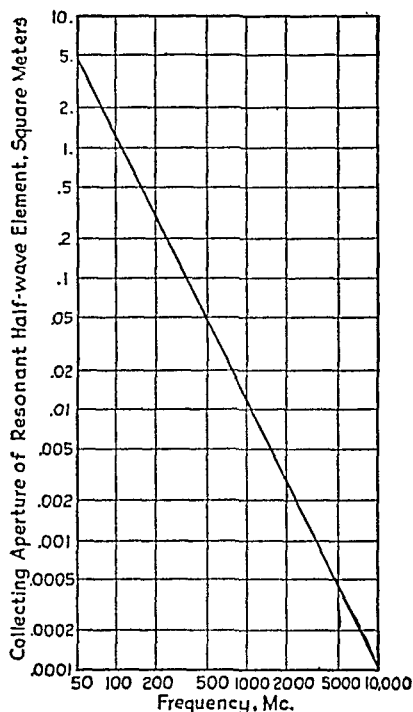


FIG. 10-8.—Collecting aperture in square meters for resonant half-wave element as a function of frequency.

The effective apertures of two direction-finder antennas over a frequency range are presented later in this chapter. The effective apertures were obtained by power-gain measurements with respect to a resonant half-wave element and the application of Eq. (10-12).

10-13. Directivity Ratio.—In Sec. 10-6 the trend in directivity of antennas is discussed as a function of aperture, and the relation of the directivity to the beam width is shown. The concept of beam width can be extended in the three-dimensional case to one of beam area. For a

transmitting antenna the *beam area* can be defined as the solid angle in square degrees through which all the radiated power would pass, if the power per unit solid angle equaled that in the direction of maximum radiation. The ratio of the number of square degrees in a sphere (41,253 sq deg) to the beam area of the antenna in square degrees can be defined as the *directivity ratio* (or simply *directivity*) of the antenna with respect to an isotropic source. Dividing this value by 1.64 gives the directivity ratio D of the antenna with respect to a resonant half-wave element. Thus, the directivity ratio D with respect to a half-wave element is

$$D = \frac{41,253}{1.64 \times \text{beam area}} \quad (10-13)$$

Power gain, unlike the directivity ratio, is assumed to include loss and mismatch effects if present. Thus, in general, the power gain when the antenna is oriented for maximum response will be less than the directivity and will equal it only when there is no loss or mismatch.

The power gain is directly related to the effective aperture and the directivity to the maximum collecting aperture. Thus,

$$G = \frac{A_e}{A_\lambda} \quad (10-14)$$

and

$$D = \frac{A_e}{A_\lambda} \quad (10-15)$$

The directivity is a dimensionless ratio so that the collecting aperture is in the same units as used for the collecting aperture of the half-wave element A_λ . When the effectivity of the antenna is unity, the maximum power gain equals the directivity and the effective aperture equals the collecting aperture, i.e., when $\epsilon = 1$, $G_{\max} = D$ and $A_e = A_\lambda$.

One method by which directivity ratio of an antenna can be obtained is to compute it from the shape of its directional characteristic. In this way the effect of loss and mismatch are excluded.

10-14. Effect of Polarization on Aperture.—The foregoing definitions of effective aperture and collecting aperture apply not only to antennas with a plane polarized response, but also to antennas with circularly or elliptically polarized responses. In each case it is assumed that the polarization of the incident wave is the same as that of the response of the antenna. This assumption gives the maximum value of the collecting aperture of a given antenna. Other values of collecting aperture will be obtained if the polarizations of the incident wave and the antenna response differ. For example, the aperture of an antenna with circularly polarized response for plane-polarized waves is one-half its aperture for circularly polarized waves, provided the direction of rotation is the same.

With the same assumptions the definitions given for power gain and directivity apply not only to antennas with a plane-polarized response, but also to those with circularly or elliptically polarized response.

10-15. Example of the Relation of the Apertures.—Consider a horn antenna having a plane-polarized response with a mouth of 2 sq m area. The physical aperture is thus 2 sq m. Assume that the power gain of the horn is measured by comparison with a resonant half-wave element and found to be 200. Suppose that the frequency is 1500 Mc (wavelength 0.2 m). The collecting aperture A_h of the half-wave element is then $0.2^2 \times 0.13 = 0.0052$ sq. m. From Eq. (10-12) the effective aperture A_e is

$$A_e = A_h G = 0.0052 \times 200 = 1.04 \quad \text{sq m}$$

From Eq. (10-11)

$$A_e = \epsilon \kappa A_p \quad \text{or} \quad \epsilon \kappa = \frac{A_e}{A_p} = \frac{1.04}{2.00} = 0.52$$

The effectivity-collectivity product for the horn is, thus, 0.52.

Next assume that the directional characteristic of the horn is measured and the beam area found to be 80 sq. deg. From Eq. (10-13) the directivity D with respect to a half-wave element is

$$D = \frac{41,253}{1.64 \times 80} = 315$$

From Eq. (10-15) the collecting aperture A_c is

$$A_c = A_h D = 0.0052 \times 315 = 1.64 \quad \text{m}^2$$

We can now tabulate the three apertures of the horn as follows:

Physical aperture A_p	= 2.00	m ²
Collecting aperture A_c	= 1.64	m ²
Effective aperture A_e	= 1.04	m ²

From these apertures the collectivity and effectivity can be obtained.

$$\text{Collectivity } \kappa = \frac{A_c}{A_p} = \frac{1.64}{2.00} = 0.82$$

$$\text{Effectivity } \epsilon = \frac{A_e}{A_c} = \frac{1.04}{1.64} = 0.63$$

The collectivity-effectivity product is

$$\kappa \epsilon = 0.82 \times 0.63 = 0.52$$

If the wave arriving at the mouth of the horn has a power of 1 watt per sq m, it then follows that the power delivered to the terminating impedance is equivalent to that gathered over an area of 1.04 sq m, or the

effective aperture. This power amounts to 1.04 watts. If there were no loss or mismatch, the power delivered would be 1.64 watts, corresponding to a collecting area of 1.64 sq. m, or the collecting aperture.

10-16. Effect of Transmission Line on Effective Aperture.—As mentioned above, the effective aperture of an antenna is affected by the value of its terminating impedance. Actually the situation is not so simple as it appears in Fig. 10-7. It is usually desirable to separate the antenna and its associated receiver physically and to connect them with some form of transmission line. The general arrangement is shown in Fig. 10-9. Some of the energy received by the antenna may be reflected at the antenna-transmission line junction *A* if an impedance mismatch is

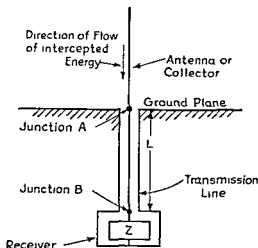


Fig. 10-9.—Collector, transmission line, and receiver showing junctions where reflections may occur.

present. Likewise some of the energy that passes *A* into the line may be reflected at *B* if there is a mismatch into the receiver impedance *Z*. When the reflections at *A* and *B* are zero, maximum energy is delivered to the receiver. If the energy reflected at *B* is not zero, there will be a standing voltage wave on the transmission line even though the reflection at *A* is zero. In general, the fraction of the intercepted energy passing *A* into the transmission line depends not only on the antenna and line impedances, but also on the impedance match at *B* and the length of line *L*. This is due to the fact that an impedance mismatch at *B* makes the effective impedance at *A* dependent on the line length *L*.

10-17. Band Width.—The band width of an antenna may be defined simply as its useful frequency range. In the case of direction-finder antennas the limits may be quite arbitrary. However, the lower limit of the range is usually the frequency at which the directional characteristic of the antenna becomes too broad for usable directional indications,

or the effective aperture becomes too small, or both. The upper limit is usually the frequency at which the directional characteristic or pattern of the antenna can no longer be interpreted readily. This commonly results when the pattern develops minor lobes of such number and size as to make a directional indication difficult or confusing.

The effective aperture of an antenna and its band width are incompatible factors. In general, the band width of an antenna is obtained at the expense of its effective aperture, or vice versa. These comments apply to fixed untuned antennas. In case the antenna is tuned to resonance by variable impedances or is adjusted physically, then the frequency band, although narrow at any one adjustment, may be moved over a relatively wide range.

Antennas described in this chapter cover a frequency range from 50 to 12,000 Mc. It has been found convenient to cover this range with four antennas with a considerable overlap in frequency between some of them. A fifth antenna is also described (half-wave elements and Adcock array) which duplicates the coverage of part of the range. Each of the antennas illustrates a distinctly different type of design. The band widths covered by the different types vary from a 1.6-to-1 frequency range to a 14-to-1 frequency range. The 1.6-to-1 frequency range corresponds to less than one octave, while the 14-to-1 range corresponds to nearly four octaves. In the design of these antennas there is usually a compromise between band width on the one hand, and effective aperture on the other.

10-18. Polarization Discrimination and Response Ratios.—As pointed out in Chap. 9, general-purpose u-h-f direction finders must be capable of taking bearings on waves that are either vertically or horizontally polarized. Several of the antennas described in this chapter fulfill this requirement by incorporating two antennas on a single rotatable base, one antenna being oriented so that the element is vertical, and the other so that the element is horizontal. By means of an r-f switch or relay, either antenna may be connected to the receiver.

To facilitate the choice of the proper antenna when a wave of unknown polarization is received, it is desirable that each antenna discriminate against waves of the opposite polarization. For example, the vertical antenna should be much more responsive to a vertically polarized wave than to a horizontally polarized wave of the same intensity. The ratio of the antenna response to two waves of opposite polarization, but of the same magnitude, may be called the *polarization-discrimination ratio*.

To illustrate, let us consider a rotatable directional vertical antenna. Assume that a vertically polarized wave of power per unit area, W , at the antenna produces a terminal power V , when the antenna is oriented for maximum response. Assume, further, that a horizontally polarized wave of the same power per unit area at the antenna produces a terminal

power V_h when the vertical antenna is again rotated for maximum response. This orientation may not be the same as for the vertically polarized wave. Then the polarization-discrimination ratio D_v of the vertical antenna is

$$D_v = \frac{V_v}{V_h} \quad (10-16)$$

In a similar way the polarization-discrimination ratio of a rotatable directional horizontal antenna may be taken as

$$D_h = \frac{H_h}{H_v} \quad (10-17)$$

Consider next both vertical and horizontal antennas on a single rotatable base. We may define what can be called the *response ratio* of the two antennas to a given polarization. The response ratio of the vertical and horizontal antennas to vertically polarized waves can be written as

$$R_v = \frac{V_v}{H_v} = \frac{V_v H_h}{H_h H_v} = \frac{V_v D_h}{H_h} \quad (10-18)$$

Similarly, the response ratio of the two antennas to horizontally polarized waves is

$$R_h = \frac{H_h}{V_h} = \frac{H_h V_v}{V_v V_h} = \frac{H_h D_v}{V_v} \quad (10-19)$$

The response ratio is particularly significant when both a vertical and a horizontal directional antenna are mounted on the same rotatable

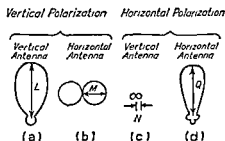


FIG. 10-10.—Typical set of patterns for a spinner with vertical and horizontal antennas showing the values used in calculating the polarization-discrimination ratio and the response ratio.

structure and a means of instantaneously connecting either to the receiver is provided. Antennas of precisely this kind are described below and are called *combination vertical-horizontal antennas*. It is desirable that both response ratios be large, in order to facilitate the choice of the proper antenna when a wave of unknown polarization is received. It is apparent that the response ratio involves the polarization-discrimination ratio of one

antenna and the ratio of effective apertures of two antennas. In general, it is desirable that the effective apertures of the vertical and horizontal antennas be nearly equal, in which case the polarization-discrimination ratio must be large in order to have a large response ratio.

To illustrate polarization-discrimination ratio and response ratio, consider the hypothetical antenna directional characteristics shown in Fig. 10-10. The pattern of the vertical antenna of a vertical-horizontal pair is shown at *a* for a vertically polarized wave of power per unit area W . The length L of the pattern maximum is assumed to be proportional to received power. The pattern of the horizontal antenna to the same vertical wave is shown at *b*. The pattern of the vertical antenna to a horizontally polarized wave of power per unit area W is shown at *c*, while the pattern of the horizontal antenna to the same horizontally polarized wave is shown at *d*.

The polarization-discrimination ratio of the vertical antenna is then

$$D_v = \frac{L}{N} \quad (10-20)$$

and of the horizontal antenna

$$D_h = \frac{Q}{M} \quad (10-21)$$

The response ratio for vertical polarization is

$$R_v = \frac{L}{M} \quad (10-22)$$

and for horizontal polarization is

$$R_h = \frac{Q}{N} \quad (10-23)$$

10-19. Bearing Deviation.—The applications to which a direction finder can be put depend to a considerable extent on the accuracy with which bearings of distant sources can be determined. As indicated in Chap. 9, reflecting objects may distort the phase front of the wave from the distant source. In general, the larger the aperture of the direction-finder antenna or collector, the smaller will be the indicated deviation due to these phase-front distortions. If the position of the reflecting objects and the direction of energy arrival in the vertical plane remain fixed with respect to the direction-finder antenna, then the bearing deviation will remain constant for a source at a given azimuthal angle. On this basis the true and observed bearings of a source can be compared for all horizontal directions and a curve of bearing deviation vs. observed bearing prepared. This is commonly called a *calibration curve*. In general, such a curve applies to a single frequency. Many such curves must be taken to determine the deviations over a frequency range.

Deviations from the true direction may also be introduced by misalignment between the direction-finder antenna and the indicating system or may be caused by the sweep circuit of the indicator. These deviations can be grouped under what may be called *instrument error*.

Another source of deviation is in the antenna itself. In the case of a directional antenna, the angle that the main lobe of the radiation characteristic makes with a reference line on the antenna structure may vary with frequency. For convenience this is called the antenna *bearing shift*. In order to determine the magnitude of this bearing shift the true and observed bearings of a source in a fixed direction may be compared over the frequency range of the antenna and under conditions substantially free from disturbances due to reflections. In the case of the combination vertical-horizontal direction-finder antennas this measurement is made for both vertical and horizontal polarization. In an antenna of symmetrical mechanical construction the measured deviations are usually less than a few degrees. When a wave polarized at 45 deg is incident on the antenna, either the vertical or horizontal element can be used for direction finding since the 45-deg polarized wave consists of equal horizontally polarized and vertically polarized components. The presence of both components tends to increase the deviation, particularly if the polarization discrimination of the antenna is low. However, if the polarization-discrimination ratio of the antenna is high, then the deviation observed on waves polarized at 45 deg is small.

10-20. Application of Direction-finder Antennas.—A directional antenna or collector can be utilized for direction finding by several methods:

1. The antenna can be rotated physically by manual or motor-driven means.
2. The antennas may be fixed physically in pairs or groups and the outputs so coupled that in effect the directional characteristic is electrically rotated at manual or motor-driven speeds.
3. The antennas may be fixed physically in pairs or groups and the outputs so coupled that the radiation characteristic is electrically rotated at radio frequency.

A number of motor-driven rotatable direction-finder antennas, or *spinners*, will be described. These antennas are used by Method 1. In addition, an example of a manually rotated and also sector-sweeping half-wave element and Adcock array is presented. The indicators associated with these antennas are described in Chap 11.

Some examples of antennas and systems employing Method 2 will be given. It might also be mentioned that the fixed crossed loops of a Bellini-Tosi direction finder¹ belong to Method 2. No antennas using Method 3 are discussed. However, a simple example of such an arrangement can be briefly stated. Consider that two horizontal half-wave elements are mounted at right angles and energized in time quadrature

¹ KEEN, R., "Wireless Direction Finding," 3d ed., Chap. 5, pp. 138-183, Iliffe and Sons Ltd., London, 1938.

This arrangement is commonly known as a *turnstile antenna*.¹ The directional characteristic of this antenna, which is a figure of eight in the horizontal plane, rotates at a rate equal to the applied radio frequency. Thus, if the elements are energized at 100 Mc, the pattern rotates 100 million times per second. By comparing the phase of the energy received by the turnstile and by a nondirectional antenna mounted above or below the center of the turnstile, the direction of energy arrival in azimuth may be deduced.

10-21. Half-wave Element and Adcock Array.—At a frequency of 150 Mc a half-wave is 1 m long (about 40 in.). A half-wave linear element for frequencies of this order may be of convenient size in some applications. When it is mounted horizontally and rotated manually, bearings can be taken on horizontally polarized sources by rotating the antenna for minimum response. There is a 180-deg ambiguity, since the directional pattern of the half-wave element has two minimums, as shown in Fig. 10-5. For aircraft applications the bidirectional pattern is not considered objectionable since the 180-deg ambiguity is usually resolved rapidly owing to the speed of an aircraft with respect to a fixed source. The fact that the individual bearings each include a 180-deg ambiguity does not influence the result, since two bearings or "cuts" intersect at only one point.

By adding an Adcock antenna consisting of two vertical half-wave elements spaced about $\frac{1}{2}$ wavelength apart at right angles to the horizontal half-wave element, the same array can be used to take bearings on either vertically or horizontally polarized sources. The resulting combination antenna may be called a *half-wave element and Adcock array*. The directions of the nulls of the Adcock in the horizontal plane coincide with those of the horizontal element, so that no bearing shift results when the source changes from vertical to horizontal polarization and reception is shifted from the Adcock to the horizontal element.

A direction-finder antenna of this kind developed in the Radio Research Laboratory found wide application. The general arrangement

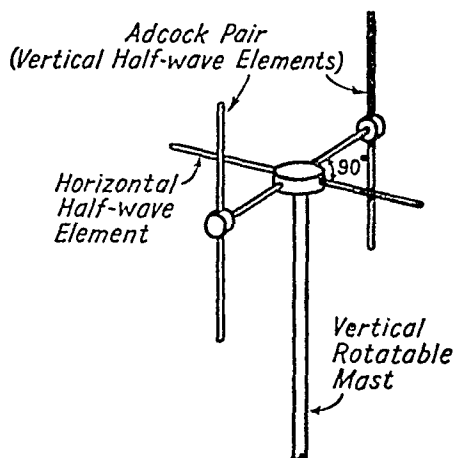


FIG. 10-11.—Direction-finder array of horizontal half-wave element and vertical Adcock antenna.

¹BROWN, G. H., A Turnstile Antenna for Use at Ultra-high Frequencies, *Electronics*, 9, 14 (1936).

of the antenna is shown in Fig. 10-11. Both the horizontal half-wave element and the vertical Adcock are rigidly supported from a rotatable mast. Installations of this array below the fuselage of an aircraft were, in general, found to give more accurate bearings than those above.

Electrical connection to the antennas is by means of twin-coaxial transmission lines. A balun (balance-to-unbalance transformer) couples each twin-coaxial cable to a single coaxial line. A line extends from each balun to a switch by which either antenna can be connected to the receiver. There is no rotating joint in the transmission line, so that the line twists as the antenna is turned. Stops are provided to prevent continuous rotation through more than about 360 deg.

The half-wave elements of the half-wave horizontal element and Adcock antenna are slender, with a length-to-thickness ratio of the order of 40. The band width of the antennas is accordingly not large, being about 1.6 to 1 for satisfactory power gain and pattern. In order to cover a wider frequency range the horizontal element and Adcock array are made as a unit that can be detached from the supporting mast and replaced by another unit or array with elements of different length and, thus, suited for operation over a different band of frequencies. In this way the frequency band 70 to 750 Mc is covered by four units, or arrays.

The antenna array is rotated by a manually operated hydraulic drive with a synchro position indicator. The receiver is first switched between the horizontal element and vertical Adcock to determine the polarization of the received wave. The antenna corresponding to the polarization is then used while the array is rotated for a minimum response. A later development has an electric motor drive, controlled by a right-left switch. In a further modification the antenna array oscillates back and forth through an azimuthal angle manually adjustable from 0 to 30 deg at a period of the order of 1 sec. This oscillating, or "sector-sweep," drive is electrically operated. With this arrangement the sector of sweep is turned slowly until the received signal strength appears to be equal on both sides of the minimum. The center of the angular sector then coincides with the direction of the source. This arrangement is described in more detail in Chap. 11.

10-22. Double-loop Spinner.—This spinner was designed for a frequency band of 50 to 300 Mc, primarily for use on aircraft where small weight, size, and wind drag are important considerations. The consequent restriction in size of the antenna to about 0.5 m places it in Group 1 of Fig. 10-6, since at 50 Mc 0.5 m amounts to about 0.08 wavelength. A loop is, thus, a logical possibility. A further requirement that the antenna operate only on horizontally polarized waves suggested the double-opposed-loop arrangement shown in Fig. 10-12a. Assume that the loops are immersed in a horizontally polarized field with the plane of

polarization parallel to the plane of the page. The directional characteristic of the loops is then as shown by the dashed lines. Ideally, the response of the loops to a vertically polarized source in the plane of the page is zero, which is an advantage as it provides a large polarization-discrimination ratio.

Carrying the arrangement of Fig. 10-12a a step further, we obtain the configuration of Fig. 10-12b. In order to connect balanced loops of this type to a coaxial or unbalanced transmission line, still another step is required. This is shown in Fig. 10-12c. Here each loop is connected into a coaxial line at points M and N , with the two lines connected in parallel at Q . The instantaneous current distribution on the loops is as shown by the arrows. The coupling from loop to coaxial line is balanced, since the current flowing from the inner conductor onto branch

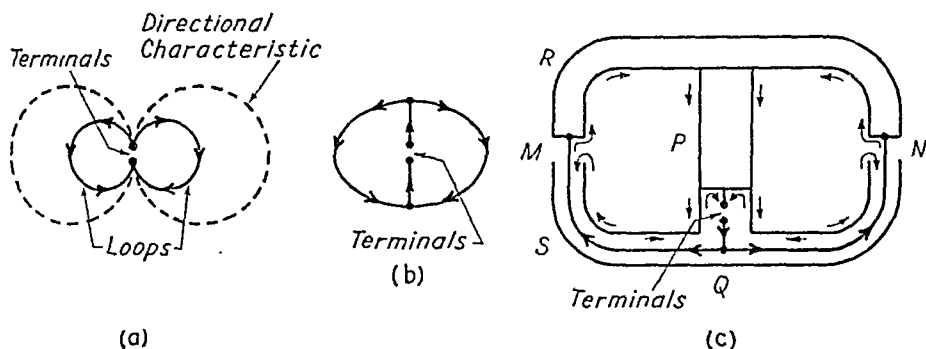


FIG. 10-12.—Steps in evolution of double-loop spinner. Current distribution shown by arrows.

R at M has an impedance path to point P which is equal to that presented by branch S to current entering or leaving the inside of the outer conductor at M . One may consider that P is at a point of electrical symmetry. Hence, a supporting column mounted perpendicular to the loops at P , should not affect the current distribution.

The double-loop spinner thus evolved operates on horizontally polarized waves and gives a bidirectional pattern. The loops are tuned to resonance by a variable capacitor in the central member of the structure, the capacitor being operated from a remote point by means of a synchro link. By means of this tuning arrangement, the antenna is made to cover a frequency range of about 50 to 300 Mc with usable power gain. The bidirectional pattern is not considered objectionable for the application intended, since the 180-deg ambiguity is usually resolved readily as explained above.

The construction of the spinner is shown in Figs. 10-13 and 10-14. The double loop is mounted on a vertical column which, together with the center member of the loop, forms a T-shaped structure. The center

element of the loop structure contains a variable air capacitor that tunes the loop to resonance, inductance coils, a frequency-range change relay, etc. The variable condenser is tuned remotely through a pair of synchros, one of which is mounted in the vertical column. Slip rings deliver current to the synchro and relay. An electric motor drive rotates the spinner at speeds of 100 to 300 rpm. The r-f transmission line from the double loop to the receiver has a rotating joint. A description of a rotating joint is given in Sec. 10-35.

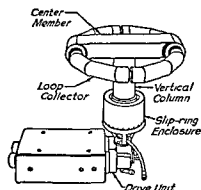


FIG. 10-13 — Double-loop spinner with drive unit.

Since the antenna is designed to respond only to horizontal polarization, the vertical column is made small in diameter to minimize the response to vertical polarization. Some response to vertical polarization

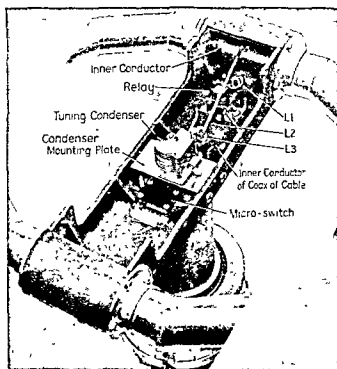


FIG. 10-14.—Center member of double-loop spinner with cover removed to show tuning components.

tion can be observed, but it is much below the response to horizontal and is generally unsymmetrical.

In the double-loop antenna the two loops are connected in opposition so that the radiation pattern is a figure of eight with minimums in the direction of the ends of the central horizontal member and maximums out from the ends of the loops. As can be seen from Fig. 10-15, the loops and the inductances L_2 and L_3 (and L_1 when the relay switch is closed) form a tuned circuit with the variable tuning capacitor. The setting of the capacitor determines the frequency to which the circuit is tuned.

Tuning the antenna to a particular frequency maximizes the current on the loop collectors at that frequency. The analysis of the complete circuit given below reveals that the circuit will be tuned at more than one frequency for a given setting of the tuning capacitor, and this is con-

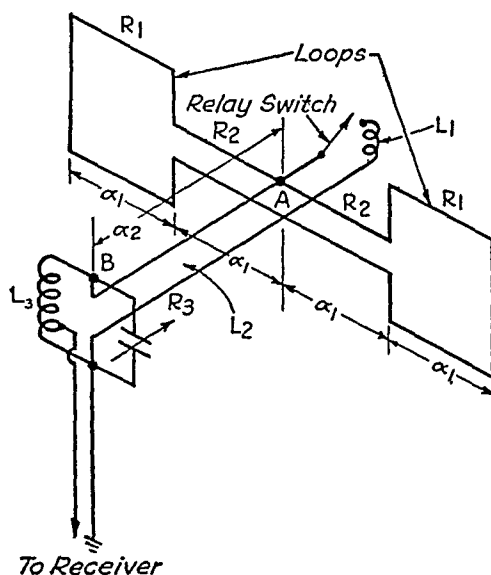


FIG. 10-15.—Equivalent circuit of double-loop spinner.

firmed experimentally. Closing the frequency-range change relay throws inductance in parallel with that already in the circuit, thus changing the frequency range covered by the tuning capacitor. With judicious choice of inductances, a fairly wide frequency range can be covered in four steps. Two of these are covered by the variable capacitor with the relay unenergized. Two others extend the frequency range at one end and fill in the gap between the first two.

10-23. Circuit Analysis of the Double-loop Spinner.—To understand the tuning of the double-loop spinner, an approximate analysis of the tuning circuit is helpful. The equivalent circuit is shown in Fig. 10-15. The loops are treated as short-circuited sections of two wire lines of characteristic resistance R_1 . The coaxial line between M and Q of Fig. 10-12 is represented by a line of characteristic resistance R_2 . The mutual imped-

ance of the loops is ignored. The electrical lengths of the loops and of the coaxial lines are taken equal to each other and represented by α_1 . With these assumptions the impedance due to one loop at point A is

$$Z_L = jR_2 \frac{(R_2 + R_1) \tan \alpha_1}{R_2 - R_1 \tan^2 \alpha_1} \quad (10-24)$$

Since the two loops are in parallel, the total impedance at point A due to both is $Z_L/2$. In Fig. 10-16 a family of curves representing values of Z_L/R_2 for different ratios of R_2/R_1 is plotted as a function of α_1 in electrical degrees. If the tuning is done at point A , it is apparent that

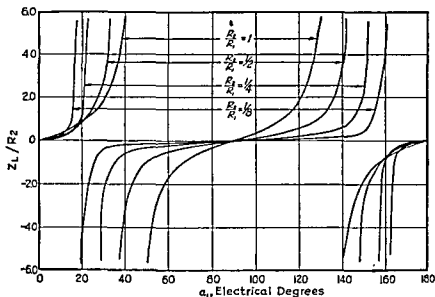


FIG. 10-16 — Ratio (Z_L/R_1) of loop impedance to characteristic resistance of coaxial line coupled to loop as a function of loop length in degrees, for several ratios of characteristic resistances of line and loop (R_2/R_1).

there will be a wide band of frequencies (where Z_L/R_2 is negative) at which capacitive tuning is impossible. It is further apparent that the larger R_2 is with respect to R_1 , the larger the band between $\alpha_1 = 0^\circ$ and $\alpha_1 = 90^\circ$ over which capacitive tuning is possible. Since this range is likely to be the one of greatest interest, one might conclude that the ratio of R_2 to R_1 should be as large as possible.

Two other factors, however, enter the picture, one being the ratio of radiation resistance to loss resistance. Because of its small size with respect to wavelength, the double-loop antenna has a very small radiation resistance. By making the loop collectors large in cross section, the loss resistance can be made small with respect to the radiation resistance. In the section of coaxial line, MQ of Fig. 10-12, most of the loss resistance

is in the inner conductor. To minimize this loss, the size of the inner conductor should be relatively large, with the result that the ratio of R_2 to R_1 may be small. The second factor affecting the size of the ratio is the slope of the impedance curve. If the ratio is small, the slope is very small over a large angular range, and hence the tuning will be broader. For this reason a small ratio is to be preferred.

Typical values of the characteristic resistances and electrical lengths are R_1 , 300 ohms; R_2 , 50 ohms; R_3 , 200 ohms; α_1 , 15 deg at 50 Mc; α_2 , $\frac{1}{3}\alpha_1$. In Fig. 10-17, $Z_L/2$ is plotted vs. frequency for these values of R_1 and R_2 . The dashed curve shows the impedance at point B of Fig.

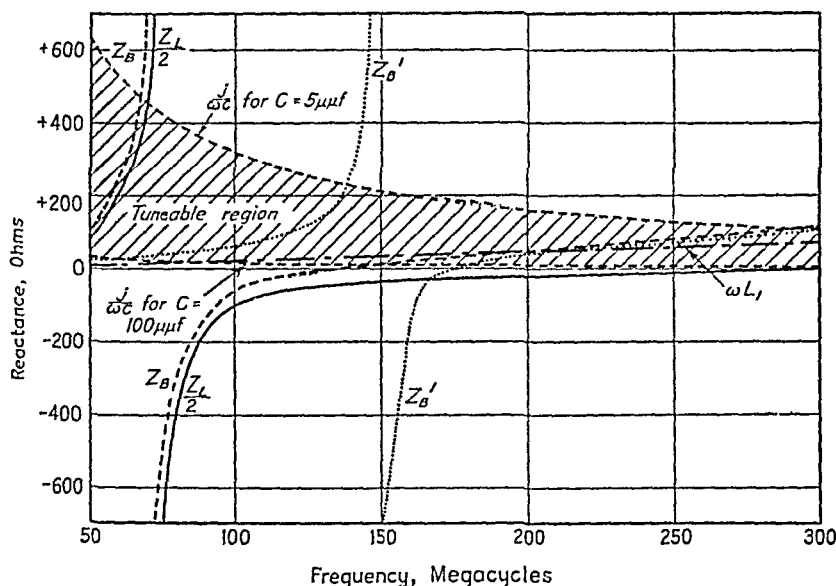


FIG. 10-17.—Various impedances involved in tuning double-loop spinner as functions of frequency.

10-15, looking toward A . The effect of the short section (α_2) of 200-ohm line (R_3) has been to raise the impedance curve so that a considerable part of it now comes within the range tunable by capacitance above 150 Mc. To fill the gap in the tuning range between about 70 and 150 Mc, parallel inductance L_1 can be added to the circuit near A . By proper choice of L_1 the impedance Z_B' as seen at B can be adjusted to permit tuning in the gap appearing above as shown in Fig. 10-17. This inductance will also extend the frequency range at its upper end by lowering the curve sufficiently to permit tuning by a 5- to 100- μf variable air capacitor. Thus by putting L_1 into the circuit, the gap in the tuning range can be filled and the range extended at its upper end. Curves of $j/\omega C$ for values of the capacitance of the air capacitor equal to 5 and to 100 μf are plotted on the graph. For values of reactance falling between

these curves, the circuit can be resonated. The inductance L_2 is large enough not to affect appreciably the impedance at B . It is incorporated in the circuit to provide an impedance transformer for coupling to the 50-ohm line leading to the receiver. This treatment of the tuning problem works out well at the lower frequencies, but deviations from the predicted behavior increase with increasing frequency.

10-24. Pattern and Power Gain Obtained with Double-loop Spinner.—Figure 10-18 is a photograph of the pattern appearing on the oscilloscope screen of one of the indicators described in the next chapter, when the double-

loop spinner is used as a component of the direction-finder system. The polarization of the signal is horizontal, and its modulation is sine wave at about 400 cycles. Bisection of the pattern through the maximums gives the bearing of the source.

The power gains obtainable with the double loop described (16 in. in diameter, 13 in. across the center member) vary from about 20 db below a resonant free-space half-wave element at 50 Mc to about 2 db below in the vicinity of 300 Mc. These measurements are made with the double-loop antenna in free space. A horizontal ground plane 10 or 12 in. from the loop, as might be present in an aircraft installation, reduces the power gain by 5 to 10 db, at angles of incidence nearly parallel to the plane. The effect is more pronounced at the low end of the frequency band.

10-25. Installation of the Double-loop Spinner.—The spinner may be mounted about 1 ft from the skin of the aircraft within a streamlined plastic housing on the forward underside or "chin" of the aircraft fuselage, as shown in Fig. 10-19a. With this location, the sensitivity and indications should be satisfactory in all directions except for a sector aft. In order to increase the loop sensitivity at the



FIG 10-18.—Double-loop-spinner pattern as it appears on indicator.

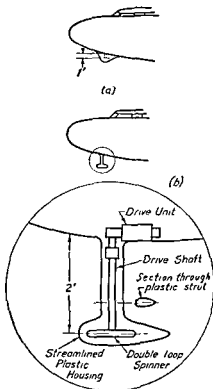


FIG. 10-19.—Installation of double-loop spinner on aircraft, showing two methods of mounting.

lower frequencies, however, where the proximity of a ground plane is particularly objectionable, a mounting at a greater distance, such as 2 ft, from the skin of the aircraft may be desirable. The greater distance can be achieved by increasing the length of the vertical column and enclosing both it and the loop in a streamlined plastic housing, as shown in Fig. 10-19b.

10-26. Combination Vertical-horizontal Corner-reflector Spinner.—

At 100 Mc a length of 2 ft corresponds to 0.2 wavelength and at 1000 Mc to 2 wavelengths. An antenna with aperture dimensions in this range falls in Group 2 of Fig. 10-6. In this range, sheet reflectors can be used to advantage. In fact it is possible to use a single sheet in the dual capacity of a reflector for both a vertical and a horizontal antenna by

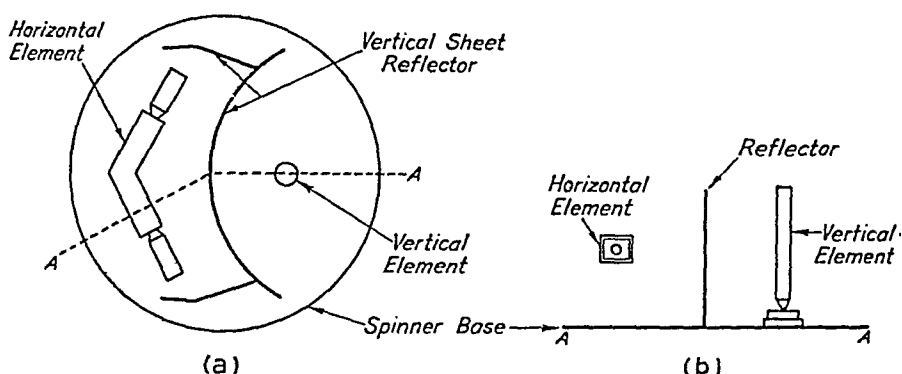


FIG. 10-20.—Plan and elevation views of vertical element and horizontal sleeve element separated by single-sheet reflector.

placing the sheet between them. Figure 10-20 shows an antenna working on this principle. A vertical element is placed on one side of a sheet and a horizontal sleeve element on the other. Either element can be connected to the receiver. Elements and reflector are mounted on a circular base, and the entire assembly forms a spinner that can be rotated.

An antenna of the same general type but covering a broader frequency band is described in this section. Instead of a reflector of relatively simple shape as in the antenna of Fig. 10-20, a number of additional reflecting surfaces are added to both sides to increase the directivity and band width of the vertical and horizontal antennas. The general arrangement is illustrated by the horizontal-plane cross section in Fig. 10-21a. A vertical element is mounted in a corner reflector of 100-deg internal angle. The corner is truncated by a surface EF , which with the sides of the corner makes the reflector walls approximate a parabolic contour. The parabolic approximation is of consequence only at the higher frequencies, where the difference in path length of energy reflected from the apex of the corner with and without the sheet EF becomes a

substantial fraction of the wavelength. Two horizontal elements in the form of a V are placed parallel to the back surfaces of the reflector. Surfaces at a 45-deg angle in the vertical plane are also placed parallel to the horizontal elements so that the reflector for the horizontal elements has the form of a truncated 90-deg corner, as shown in Fig. 10-21b.

Both antennas are mounted on a rotatable circular base forming a combination vertical-horizontal corner-reflector spinner, or simply a *corner spinner*. Either the vertical or horizontal elements may be connected to a receiver by means of a relay-operated single-pole double-throw switch in the r-f transmission line. For convenience this is referred

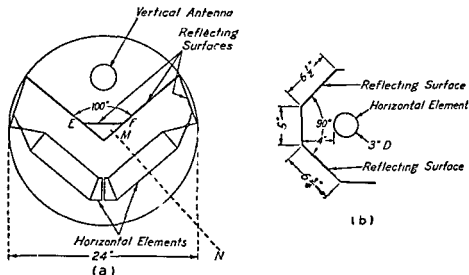


FIG. 10-21.—(a) Plan and (b) elevation section along MN of corner spinner.

to as an r-f relay. It is mounted on the spinner under the reflecting surfaces. A typical relay is described in Sec. 10-37.

The corner spinner thus consists of one antenna responsive to vertically polarized waves and another responsive to horizontally polarized waves, both mounted back-to-back on a rotatable base. Since the directional characteristics point in opposite directions, a 180-deg bearing shift would be produced when the receiver is switched from the vertical to the horizontal antenna, or vice versa, if compensation were not provided. This is accomplished by means of another relay-operated switch located in the indicator unit and working simultaneously with the one in the r-f transmission line. The second switch reverses connections to the indicator cathode-ray tube and thereby reverses the observed pattern direction.

The horizontal antenna of the corner spinner is shown in Fig. 10-22. A plastic radome forming a weatherproof enclosure for the antenna has

been removed from the base in this figure. The r-f relay is mounted on the spinner base-plate behind one of the reflecting surfaces below the horizontal element.

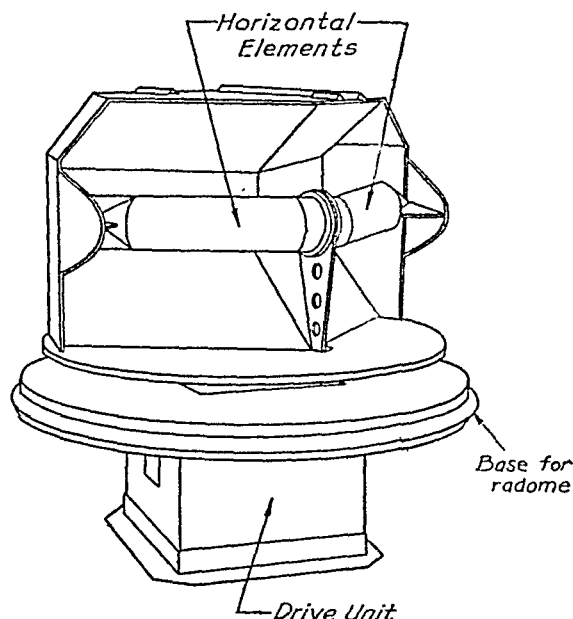


FIG. 10-22.—Horizontal antenna of corner spinner.

The spinner operates over a nominal frequency range of 100 to 1400 Mc and is designed for rotation at speeds up to 300 rpm. The spinner

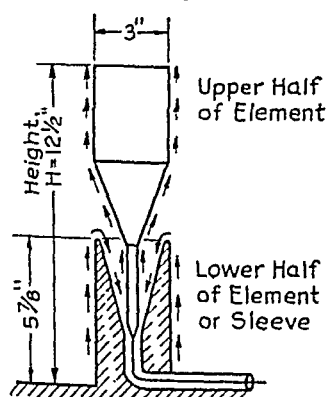


FIG. 10-23.—Cross section through vertical element of corner spinner. Arrows show instantaneous current distribution.

measures 24 in. in diameter by 15 in. high and was developed primarily for ship-borne applications. The weight of the spinner with balance weights is only about 20 lb. The reflecting surfaces of the spinner are largely of 0.010-in. stainless steel sheet, and the elements are of thin-walled brass tubing. A rotating joint, such as described in Sec. 10-35, is located in the transmission line from the antenna to the receiver.

10-27. Vertical Antenna of Corner-reflector Spinner.—The construction of the vertical element of the spinner is shown in cross section in Fig. 10-23. The element is a vertical sleeve collector of large cross section enclosed in, and supported by, a plastic sleeve, which does not appear in Fig. 10-23. The instantaneous current distribution on the sleeve element when the height H is less than $\frac{1}{2}$ wavelength is shown by the

arrows. This sleeve element matches a 50-ohm transmission line relatively well over a wide frequency band. Furthermore, the currents on the sleeve element that react directly with the incident wave (*i.e.*, the currents on the 3-in. diameter sections and the conical taper of the upper half) are in substantially the same phase for heights H less than 1 wavelength. As a result the maximum of the directional pattern in the vertical plane is more nearly at right angles to the sleeve element, or at a vertical angle of incidence nearer to 0 deg, than if the vertical element had no sleeve surrounding its lower half.

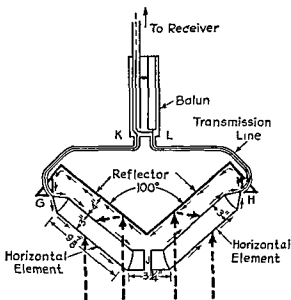


FIG. 10-24.—Horizontal elements of corner spinner. Small arrows show current distribution.

10-28. Horizontal Antenna of Corner-reflector Spinner.—The arrangement of the horizontal elements is indicated by Fig. 10-24. The elements, which are parallel to the back of the vertical-element reflector, are of large cross section and form a V of 100-deg internal angle. At frequencies for which the total length of both elements is less than 1 wavelength the instantaneous current distribution is as shown. The current flows in one direction along the horizontal elements and in the reverse direction along the surface of the reflector. The direction of energy flow from a source in front of the aperture (below in the figure) toward the antenna elements is indicated diagrammatically by the large dashed arrows.

The ends of the horizontal elements are connected to the center conductor of a coaxial cable at the points G and H . Cables of equal length extend from G to K and H to L where both cables are connected by a

balance-to-unbalance transformer, or balun, to a single coaxial cable that connects to the receiver. Some balun-design considerations are treated in Sec. 10-39. It will be noted that the ends of the horizontal elements are not connected together at point J at the apex of the V, although there is considerable capacitance at the gap. At most frequencies, closing the gap or leaving it open makes little difference in the directional characteristic of the antenna. Over a wide frequency range the open gap is, however, to be preferred. With the gap closed it is more obvious that the antenna is but a form of a horizontal loop. From another viewpoint, the horizontal elements GJ and HJ are part of a reentrant transmission line which also includes the coaxial cables to the balun.

From still another standpoint, the horizontal elements are branches of a V-antenna coupled into transmission lines at the open end of the V rather than at the apex. The fact that the connection is made at the ends rather than the apex of the V is of fundamental importance. To explain the basis for this difference, consider the two V-antennas illustrated in Fig. 10-25. For convenience the transmitting case will be considered. Energy entering the V of Fig. 10-25a at the apex progresses along the branches of the V as a traveling electromagnetic wave. If it is assumed that the branches of the V are long in terms of wavelengths, the amount of energy reflected from the open ends will be small since most of the energy is radiated before the wave reaches the open end. Under these conditions the directional characteristic of each branch is as indicated and the maximum of the total characteristic is in the direction shown by the large arrow.

Energy enters the V of Fig. 10-25b from the opposite end. It is assumed that the sections P and Q do not radiate. If the branches R and S of the V are long, then the directional characteristics of the individual branches are as shown and the maximum of the total characteristic is in the direction of the large arrow. This V-antenna is directional in a direction opposite to that of the apex-connected V merely because the energy is introduced at the opposite end of the V.

The horizontal elements of the corner spinner are of relatively large cross section. Since, as a wave travels along the element, more radiation takes place from an element of large cross section than from one of small cross section, the horizontal elements tend to be unidirectional in the way

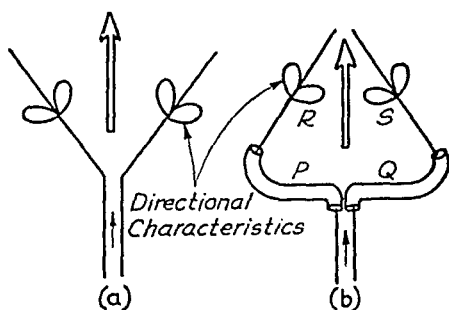


FIG. 10-25.—(a) Apex-connected V-antenna and (b) end-connected V-antenna.

described above, even though they are relatively short. By reciprocity the directional characteristic in the transmitting case is the same as in the receiving case. Thus, the direction of energy flow for maximum sensitivity of the corner spinner is as indicated by the large dashed arrows in Fig. 10-24.

10-29. Directional Characteristics.—A series of corner-spinner patterns over a frequency range from 90 to 2000 Mc is presented in Fig. 10-26. The nominal frequency range of the antenna is about 100 to 1400 Mc. The patterns at other frequencies are presented as a matter of interest. The patterns are for a vertical angle of incident energy of about 0 deg. The line drawings were made from photographs of the

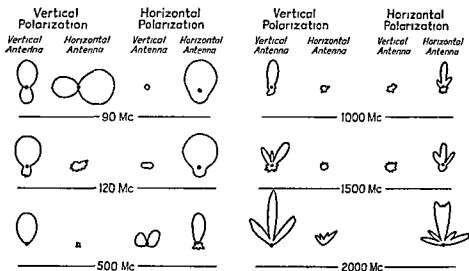


Fig. 10-26 — Corner-spinner patterns.

cathode-ray-tube screen of the direction finder with which the spinner was used. In these patterns the deflection is approximately proportional to the field intensity.

As the frequency increases above 90 Mc, the patterns tend to become sharper and the back lobes smaller. The response ratio of the antennas for both polarizations is large in all cases shown except two. At 90 Mc, with vertical polarization the horizontal antenna shows a large response with a pattern of shape similar to that of the vertical antenna but turned through 90 deg. However, at this frequency the response ratio for horizontal polarization is large so that it is possible to distinguish between vertical and horizontal polarization. At 500 Mc, with horizontal polarization, the vertical antenna shows a moderate response, but the shape of the pattern is unsymmetrical and is distinctly different from that produced by the horizontal antenna. Thus the proper antenna can readily

be selected. At most frequencies the response ratio for both polarizations is high.

It is apparent from Fig. 10-26 that the antenna patterns are characteristic of the frequency. Thus the frequency of an unknown source can be estimated by the shape of the received pattern. For example, if the pattern shown in Fig. 10-26 for the horizontal antenna with horizontal polarization at 1000 Mc is observed on the indicator, it signifies that the frequency of the source is about 1000 Mc. Suppose, however, that the receiver dial indicates a frequency of 500 Mc. From this it can be concluded at once that the second harmonic of the receiver oscillator is beating with the signal received by the antenna. Retuning the receiver to about 1000 Mc should result in stronger reception of the source. If the source frequency was 500 Mc, the pattern would appear as shown. The difference between the pattern for 500 Mc and the one for 1000 Mc is unmistakable.

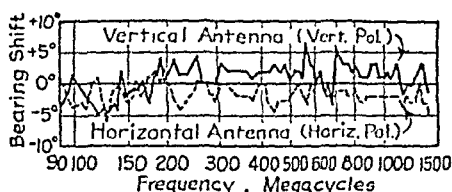


Fig. 10-27.—Bearing shift as a function of frequency for vertical and horizontal antennas of corner spinner.

10-30. Bearing Shift and Effective Aperture.—In Fig. 10-27 curves of the bearing shift (see Sec. 10-19) of the vertical and horizontal antennas of the corner spinner are shown as a function of frequency. With the exception of two points, all values fall within plus or minus 5 deg. In Fig. 10-28 the measured effective aperture in square meters of the vertical and horizontal antennas of the spinner is presented as a function of frequency. The ratio of the effective aperture to the physical aperture of the antennas (or effectivity-collectivity product) is about 0.1 for most of the nominal frequency range. This means that even though the antennas cover a very wide frequency range (over 1000 per cent) the antennas have only about 10 db less gain than they might have on the basis of their physical aperture. The values of effective aperture shown take account of losses in the cable through the driving unit and in the rotating joint.

10-31. Combination Vertical-horizontal Parabolic-reflector Spinner. At 1000 Mc, an antenna of 2 ft aperture is about 2 wavelengths across. Thus for an aperture of about 2 ft, and frequencies of 1000 Mc and above, we are in the region where antennas of Group 3 of Fig. 10-6 may be employed to advantage. A suitable direction-finder antenna belonging

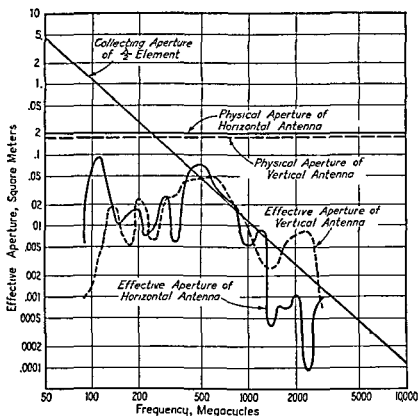


FIG. 10-28—Effective aperture in square meters of corner spinner antennas as a function of frequency. Effect of losses in r-f relay, rotating joint, and drive unit included. Physical apertures of antennas and collecting aperture of a resonant half-wave element indicated for comparison.

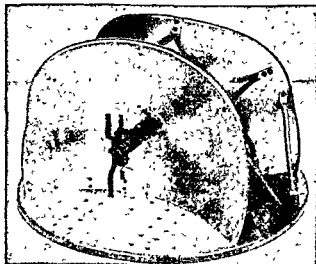


FIG. 10-29.—Photograph of double-parabolic spinner.

to this group consists of a combination of a vertical and a horizontal antenna, each with its own parabolic reflector. It may be designated as a *combination vertical-horizontal parabolic-reflector spinner*, or simply as a *double-parabolic spinner*. Figure 10-29 shows an over-all view with the side responsive to vertical polarization facing the reader. This antenna has a maximum dimension of about 20 in. and is used on frequencies from 1000 to 5000 Mc. The size and weight are sufficiently small to make the antenna attractive for airborne installations, as well as for ship and land applications. The antenna is designed for rotation at speeds of the order of 300 rpm. The weight of the airborne antenna is about 10 lb.

10-32. Paraboloid Reflectors.—The parabolic reflectors are actually parabolas of revolution, or paraboloids. The antenna was originally designed to mount inside a hemispherical radome or blister of such size that it was necessary to cut off each of the two paraboloid components (18 in. diameter, $4\frac{1}{2}$ in. focal length, $4\frac{1}{2}$ in. depth) by a plane parallel to the axis of rotation of the paraboloid. Physically this plane is formed by a flat metal disk 20 in. in diameter, on which the paraboloids are mounted in a back-to-back relation, as shown in Fig. 10-29. The paraboloids have a shape factor of 4 (ratio of diameter to focal length) and a focal length equal to the depth. Since the antenna is designed to cover a frequency band of about 5 to 1, and is of restricted size and shape, some departures from an optimum single-frequency design were necessary. At 1000 Mc the diameter of the paraboloid (18 in.) is about 1.5 wavelengths and at 5000 Mc about 7.5 wavelengths.

An r-f relay, by which either antenna can be connected to the transmission line to the receiver, is mounted between the paraboloids on the 20-in. disk. A rotating joint is located in the transmission line between the relay and the receiver. A description of the rotating joint is given in Sec. 10-35 and of the r-f relay in Sec. 10-37.

10-33. Primary Antennas.—Each paraboloid reflector has a primary antenna or collector mounted near its focus. The primary antenna in each case is a "balanced-sleeve" element. The band width and impedance characteristics of this type of antenna are discussed in Chap. 5. In one paraboloid the sleeve element is mounted vertically, making this array responsive to vertically polarized waves. In the other paraboloid, the sleeve element is horizontal and is responsive to horizontally polarized waves. In each case a reflector element is placed near the sleeve element to increase the response of the sleeve to energy reflected from the paraboloid. The arrangement of the primary antenna is illustrated in Fig. 10-30. At high frequencies the current flow on the end cylinders of the sleeve element is suppressed by means of small side stubs or rods mounted perpendicular to the end cylinders near each end of the sleeve section,

as shown in Fig. 10-29. When the over-all length of the side stubs is about $\frac{1}{2}$ wavelength or more, they present a lower impedance than do the end cylinders and thus reduce the current on the end cylinders. Near 1000 Mc, currents flow over substantially the entire length of the sleeve antenna and of the reflector. Above 3500 Mc it is necessary that the reflector be closer to the part of the sleeve element carrying most of the current in order to maintain proper directivity. This is done, in effect, by the bent shape of the reflector, which places it nearer to the central or sleeve portion. In this way, the primary antenna maintains its direction of maximum sensitivity more or less along the axis of the paraboloid, and the pattern remains sufficiently broad to accept energy arriving via most of the reflector area.

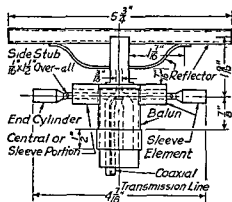


FIG. 10-30.—Sleeve element and reflector assembly of primary antenna of double-parabolic spinner

be seen in Fig. 10-29, the reflector of the production model is a single bow-shaped element. This was found to be easier to die-cast than the original design (Fig. 10-30) and gave equivalent results.

10-34. Characteristics of the Parabolic-reflector Spinner.—Patterns of the directional characteristic of the antenna in the horizontal plane are shown in Fig. 10-31. This is the plane of the 20-in.-diameter base disk. The vertical patterns are those of the vertical antenna for vertical polarization, and the horizontal patterns are those of the horizontal antenna for horizontal polarization. In these patterns the deflection is approximately proportional to the field intensity.

The maximum response of the horizontal array makes a vertical angle of incidence with the mounting plate that varies with frequency, since the plate serves to deflect the beam in elevation. The angle of maximum response is about 25 deg at 1000 Mc, 15 deg at 1100 Mc, and decreases to 10 deg at 2000 Mc, remaining relatively constant at higher frequencies.

The antenna power gain as compared with a free-space half-wave element varies over the 1000-to-5000-Mc frequency range. The gain of

Each primary antenna transfers its received energy to a 50-ohm coaxial (unbalanced) transmission line. To allow a balanced connection at the antenna, a balance-to-unbalance transformer, or balun, is provided, mounted in the sleeve antenna support (Fig. 10-30).

The various components of the sleeve and reflector assembly are identified in Fig. 10-30. As may

the horizontal array also differs from that of the vertical array. The measured effective aperture of the antenna in square meters is shown by the curves of Fig. 10-32. It is apparent that the effective antenna apertures are greater than the collecting aperture of a half-wave element at

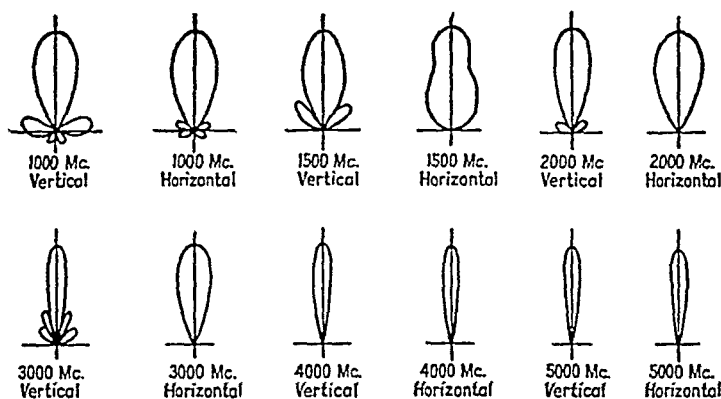


FIG. 10-31.—Double-parabolic spinner patterns.

most frequencies. However, the effective aperture of each antenna is less than the physical aperture at nearly all frequencies. The ratio of effective to physical aperture of the antennas averages about 0.1 which indicates that the antennas have about 10 db less gain than they might

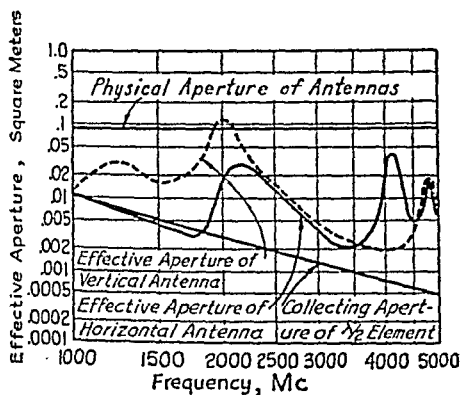


FIG. 10-32.

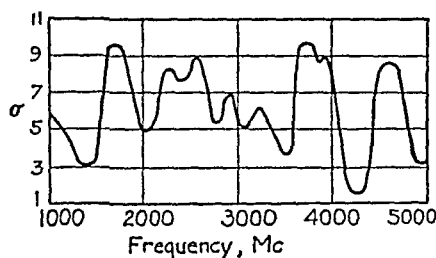


FIG. 10-33.

FIG. 10-32.—Effective aperture in square meters of double-parabolic spinner antennas as a function of frequency. Effect of loss or mismatch due to relay and rotating joint not included. Physical aperture of antennas and collecting aperture of a resonant half-wave element indicated for comparison.

FIG. 10-33.—Standing-wave-ratio σ on 50-ohm line of horizontal antenna of double-parabolic spinner.

have on the basis of their physical aperture. The standing-wave ratio on a 50-ohm line measured with the horizontal antenna operating as a transmitting array or radiator varies from less than 2:1 to almost 10:1 over the frequency range, as shown in Fig. 10-33.

The polarization-discrimination ratio of the antennas is high over the entire frequency range. The polarization-discrimination ratio of both antennas is never less than 6 db, and for the greater part of the frequency range it is at least 15 to 20 db. Since the effective aperture of both antennas is of the same order at most frequencies, the response ratio to both polarizations is usually large.

This antenna has given satisfactory direction-finding performance when mounted on the chin of an aircraft fuselage or on a ship's mast.

10-35. Spinner Auxiliaries. Rotating Joint.—In each of the three spinners described above, a coaxial transmission line is used to transfer

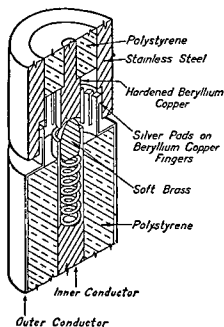


FIG. 10-34.—Cross-sectional sketch of rotating joint

of a stainless-steel tube. The fingers are stationary while the tube rotates. Contact between the stationary and rotating parts of the inner conductor are made by a soft brass tip held by a spring against a hardened beryllium-copper cup. The tip is stationary and the cup rotates.

10-36. Slip Rings.—In the section on the double-loop spinner it is mentioned that slip rings are required to carry the current for operating the synchro and relay. A satisfactory choice of materials is phosphor-bronze for the rings and silver-carbon for the brushes.

The corner spinner is also provided with slip rings and brushes for transferring the relay current. The currents handled by the slip rings in the above applications are either direct currents or power-frequency alternating currents (60 to 400 cps).

r-f energy from the spinner to the receiver. To permit continuous rotation of the spinner, a rotating joint in the transmission line is required. It is placed in the transmission line passing through the drive unit. The general construction of a coaxial rotating joint which has been used is shown in Fig. 10-34. The effective ratio of the diameters of the inner and outer conductors is maintained nearly constant through the joint. In this way but little discontinuity is introduced and r-f reflections are not serious up to frequencies of about 5000 Mc. Contact between the rotating and stationary parts of the outer conductor is made by silver pads on hardened beryllium-copper fingers that ride on the polished surface

10-37. Radio-frequency Relay.—Both the corner spinner and the double-parabolic spinner have an r-f relay for connecting either the vertical or horizontal antenna to the transmission line to the receiver.

The relay used in the double-parabolic spinner is designed for operation on frequencies up to about 5000 Mc. It is in effect a single-pole double-throw switch operated by a solenoid and lever arrangement. The relay is constructed of coaxial elements in order to maintain a characteristic impedance of 50 ohms. The relay or switch consists of a fixed center contact and two moving contacts, each mounted on a flexible leaf of phosphor-bronze, as illustrated in Fig. 10-35, which shows the relative positions of the inner conductors of the transmission lines. The leaves are actuated by a solenoid.

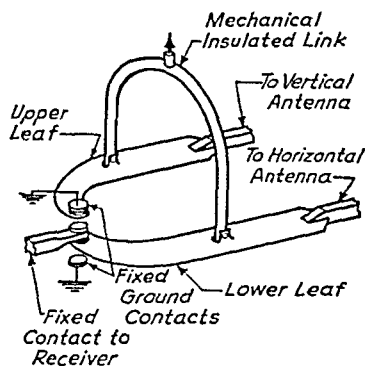


FIG. 10-35.—Sketch of r-f relay showing basic components of inner conductors. Arrangement of complete relay is shown by Fig. 10-37.

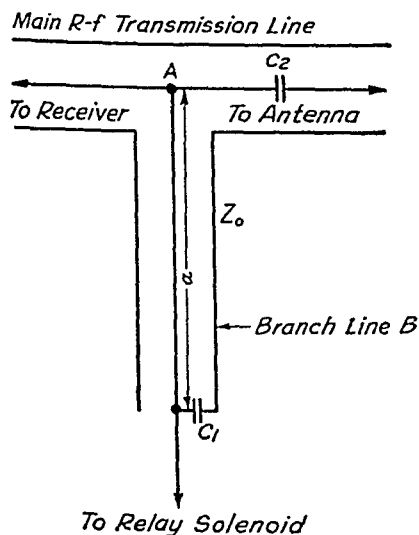


FIG. 10-36.—Schematic diagram of circuit for separating relay current from r-f transmission line.

of the two moving contacts, the ground circuit being completed to the contact not in use.

10-38. Relay-current Separator.

In some applications of the double-parabolic spinner, the current for actuating the relay solenoid is carried by the r-f coaxial line from the receiver. The components for separating the relay current from the r-f transmission line at the relay on the double-parabolic spinner are shown schematically in Fig. 10-36. A branch line *B* is connected with the main r-f transmission line at *A*. The branch itself can be considered as a length of transmission line of characteristic impedance Z_0 and electrical length α . The inner and outer conductors are joined

at the end of the branch remote from *A* through the low impedance of capacitor C_1 . If the capacitor impedance is sufficiently small, it effectively short-circuits the line so that the impedance Z of the

branch line, which appears in parallel with the main-line at *A*, can be written

$$Z = jZ_0 \tan \alpha \quad (10-25)$$

Assume, for example, that Z_0 is 200 ohms and the frequency is such that α is 45 deg, then

$$Z = j200 \tan 45^\circ = j200 \quad \text{ohms}$$

If the characteristic impedance of the main line is 50 ohms, Z will have relatively small effect at this frequency. At twice the frequency, α is

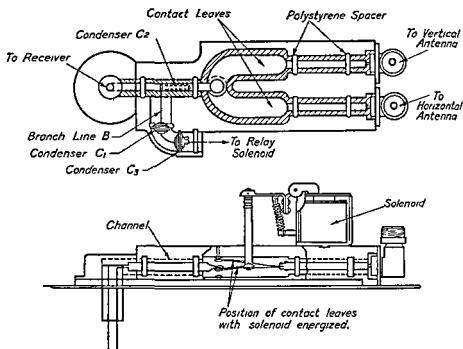


FIG. 10-37.—Cut-away views of r-f relay and relay-current separator.

90 deg and Z , on this basis, is infinite. Thus when α is 90 deg, the branch line should have no effect. When α is 180 deg, Z is zero and the branch line short-circuits the main line. If Z_0 is sufficiently large, there will thus be a range of frequencies over which the branch line has relatively little effect on the main line. This range may extend, for example, over a 5:1 ratio or over values of α from about 30 to 150 deg.

The current for the relay is taken from the branch line at the capacitor C_1 . Another capacitor C_2 is inserted in the main line to prevent the direct current or power-frequency alternating current from entering the antenna and being short-circuited to ground in the balun. The center conductor of the branch line used in the double-parabolic spinner

consists of a fine wire in order that its characteristic impedance be high. This wire, the capacitors corresponding to C_1 and C_2 , and the relay are shown in cut-away views in Fig. 10-37. Another capacitor C_3 is shown in parallel to C_1 to provide added admittance to ground at the lower radio frequencies. The standing-wave ratio produced on a 50-ohm line, terminated in its characteristic impedance, by the insertion of the relay of Fig. 10-37 with the separator is shown in Fig. 10-38, for the frequency range 1000 to 5000 Mc. The standing-wave ratio is under 2:1 below 3500 Mc and rises as high as 6:1 near 5000 Mc. A device similar to the separator described is also required at the receiver end of the r-f line to inject the direct current or power-frequency alternating current into the transmission line.

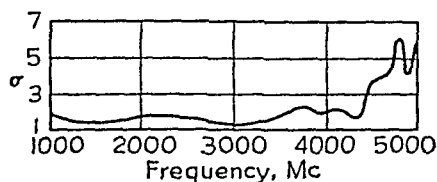


FIG. 10-38.—Standing-wave ratio σ produced by r-f relay with separator on a 50-ohm line terminated in its characteristic impedance.

10-39. Balun.—The horizontal antenna of the corner spinner has a balun (balance-to-unbalance transformer) for coupling the two r-f transmission lines from the ends of the horizontal elements to a single line to the receiver (Fig. 10-24). The primary antennas of the double-parabolic spinner also include a balun for coupling to the balanced sleeve element (Fig. 10-30).

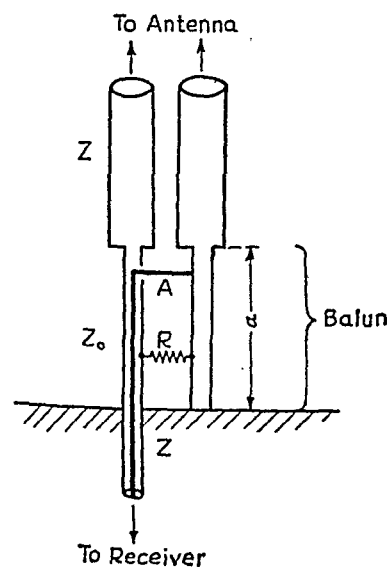


FIG. 10-39.—Schematic diagram of wide-band balun.

The balun provides a balanced transformation not only when the length is 90 deg but also for a wide range of lengths. Figure 10-39 shows a single coaxial line of characteristic impedance Z connected to a balanced two-wire line of the same characteristic impedance by means of a balun of characteristic impedance Z_0 and length α . The reasoning involved in choosing the length α and the characteristic impedance Z_0 is somewhat similar to that given above

for the branch line of the relay-current separator. The impedance of the balun line which appears across the junction of the coaxial and two-wire lines at A is

$$Z = jZ_0 \tan \alpha \quad (10-25)$$

If Z_0 is sufficiently large, a wide range of values of α will result in a balun reactance at A which has relatively small effect on the two lines of impedance Z . When α is small, the reactance appearing at A is inductive, and advantage may in fact be taken of this to balance out capacitive reactance from the antenna. Thus values of α from 20 to 160 deg may be permissible.

In the corner spinner the band width is over 14 to 1, and it is necessary to work the balun through the point where α is 180 deg. This can be accomplished by adding a resistance R approximately equal to the characteristic impedance of the balun across the balun about one-third of the way up from the grounded end, as shown in Fig. 10-39.

10-40. Circularly Polarized Horn with Parabolic Spinner.—At frequencies above 5000 Mc and for apertures of about 0.5 m we are far into the region of antennas of Group 3 of Fig. 10-6. Under these conditions direction-finder antennas of yet different design are possible. In this section a direction-finder spinner is described which operates over a nominal frequency range of 5000 to 12,000 Mc. The antenna is about 15 in. in diameter by about 29 in. in height. It differs from the types already described in that the primary antenna is a horn collector pointing upward, with circularly polarized response, which will be referred to as a *circularly polarized horn*. The horn is stationary and connects to a wave-guide transmission line. Energy is deflected into the horn by a rotating parabolic reflector mounted at a 45-deg angle above the mouth of the horn. The reflector rotates at speeds of several hundred rpm. The antenna has equal response to waves at all bearing angles in the

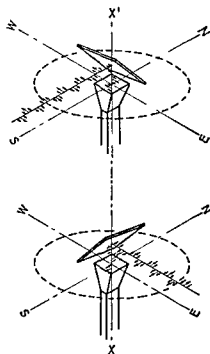


FIG. 10-40.—Circularly polarized horn and plane reflector.

horizontal plane and for all angles of polarization. The antenna is also responsive to circularly polarized waves having the proper direction of rotation. The use of waveguide transmission line permits efficient operation at much higher frequencies than is possible with coaxial line.

The double-loop spinner described in Sec. 10-22 requires a rotating r-f joint, while the corner spinner and double-parabolic spinner employ both an r-f joint and an r-f relay. In the circularly polarized parabolic-

reflector spinner discussed in this section, the need for a rotating r-f joint is eliminated by mounting the horn in a fixed position so that it points upward into a reflector, as shown in Fig. 10-40. Since the antenna responds to circular polarization, one antenna performs the task of two in the double-parabolic spinner, and the need for an r-f relay is eliminated.

The antenna pattern is rotated by causing the reflector to rotate about the axis XX' while the horn remains stationary. With this arrangement the direction of electric or \mathcal{E} lines entering the horn aperture depends upon the direction from which the wave is arriving. If, for example, a horizontally polarized wave arrives from the south the \mathcal{E} lines lie along the east-west (EW) axis. It follows that the horn must be capable of the same response regardless of the orientation of polarization of the imping-

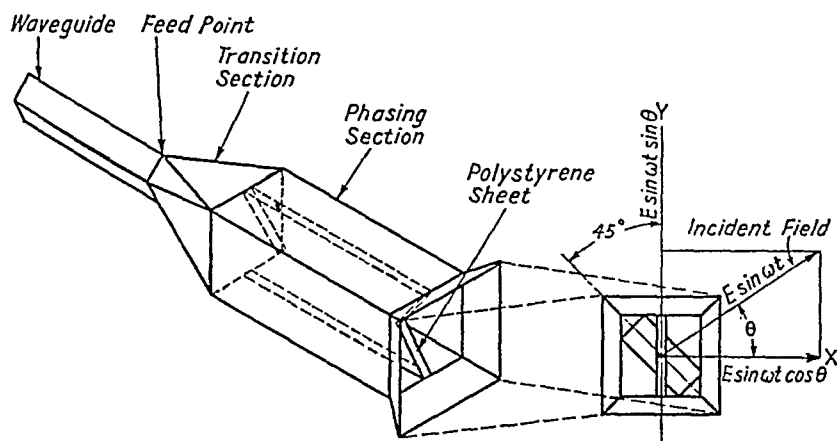


FIG. 10-41.—Circularly polarized horn.

ing wave in order to make the response of the antenna independent of the direction of arrival of the wave. Such a horn is sketched in Fig. 10-41. It consists of two essential parts, a phasing section and a transition section, which orients the phasing section at 45 deg with respect to the waveguide as shown. The theory and design of this type of horn are discussed in Chap. 6. It suffices here to say that it operates as a collector and to show that its response is independent of the orientation of impinging \mathcal{E} lines.

In the proper frequency band the cross-sectional dimensions of the horn are such that energy entering the horn will traverse it in two modes, viz, the $TE_{0,1}$ and the $TE_{1,0}$ modes. Thus the electric field at any point in the horn is the resultant of a field parallel to the polystyrene insert, or phaser, and of a field perpendicular to the insert. In general, these two fields will not be in phase. The length of the insert is so chosen that at the entrance to the waveguide the field perpendicular to the insert will lead that parallel to the insert by 90 deg, provided the two fields were in

phase at the mouth of the horn. It is impossible for this phase difference to be exactly 90 deg for all frequencies in the frequency band of the antenna, but by the proper choice of dimensions and dielectric material for the phaser a phase difference of 90 deg may be approximated closely enough for practical purposes. The dimensions of the waveguide are such that it will accept only the $TE_{0,1}$ mode in which the electric field is parallel to the smaller dimension of its rectangular cross section.

If the reflector of Fig. 10-40 is an infinitely large flat sheet, then the antenna pattern will be essentially the same as the horn pattern, as shown

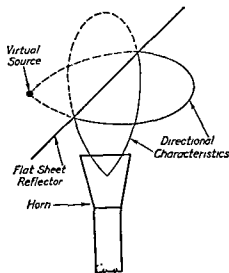


FIG. 10-42.—Diagram showing directional characteristic of horn to be unchanged by flat sheet reflector.

in Fig. 10-42. In order to keep the space patterns of the over-all antenna the same for all bearings of the reflector, the space pattern of the horn alone should be symmetrical with respect to the longitudinal axis of the horn. A horn of this type is shown in Fig. 10-41. In this horn the cross section of the phaser and aperture is square so that the patterns of the horn will nearly satisfy the required symmetry about the axis of the horn.

Another way of expressing the fact that the circularly polarized parabolic-reflector spinner responds equally well to vertical and horizontal polarization is to say that its polarization-discrimination ratio is

unity. Thus, the polarization of the received wave cannot be determined with this antenna.

The horizontal pattern of the complete antenna should be fairly sharp to permit accurate determination of direction. On the other hand, the vertical patterns should be broad so that if the antenna is mounted on a ship or airplane the signal will not disappear as the ship rolls or the airplane banks. This is achieved in the antenna pictured in Fig. 10-43 by shaping the reflector so that it forms a section of a parabolic cylinder with its longitudinal axis intersecting the longitudinal axis of the horn at an angle of 45 deg. A more complete discussion of the design of this type of reflector is presented in Chap. 6.

The antenna patterns in the horizontal plane for horizontal and vertical polarization are shown in Fig. 10-44, for frequencies of 5300 and 10,000 Mc. The directional characteristics of the antenna in the vertical plane are shown in Fig. 10-45

The spinner shown in Fig. 10-43 is constructed so that the reflector is mounted on a ring gear that revolves about the axis of the horn. This ring gear is motor-driven. A synchro is also turned at the same speed as

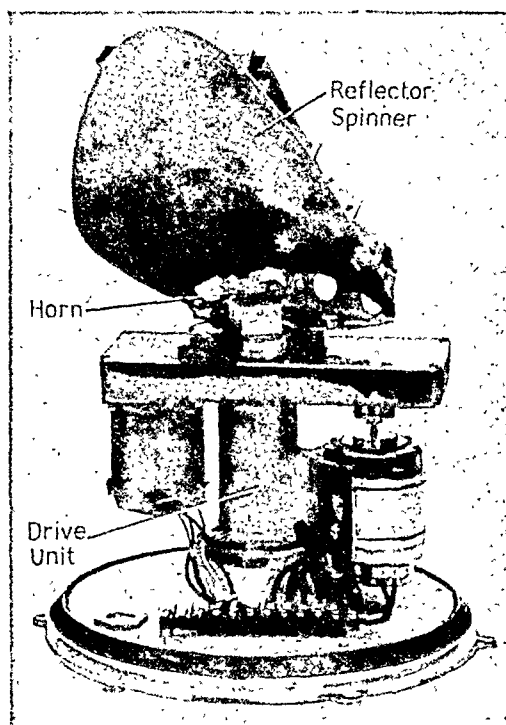


FIG. 10-43.—Photograph of circularly polarized horn and parabolic spinner with radome removed.

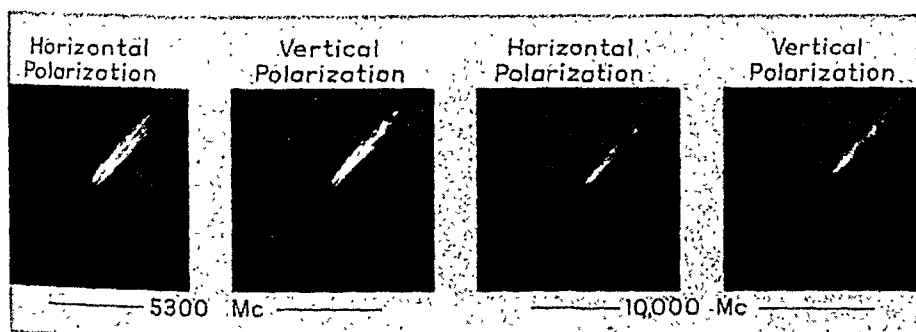


FIG. 10-44.—Photographs of patterns of circularly polarized horn with parabolic spinner as they appear on indicator. Patterns are in horizontal plane for both horizontally and vertically polarized waves. Source is at a bearing of about 45 deg.

the antenna. The synchro provides the indicator with information as to the instantaneous bearing of the reflector. This spinner drive unit is designed primarily for shipborne applications. In airborne applications

the drive unit is made as shallow as possible to facilitate installation in restricted quarters.

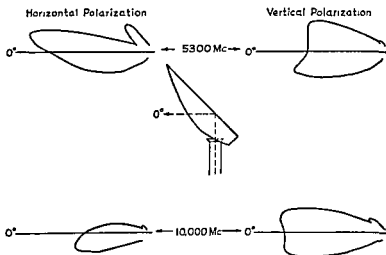


FIG. 10-45.—Patterns of circularly polarized horn with parabolic spinner in vertical plane

10-41. Fixed Antennas.—At low frequencies a loop antenna is the leading candidate for a directional antenna of a size that can be rotated

conveniently. The sensitivity of the loop may, however, be considerably less than that of a linear half-wave element, even though the loop consists of more than one turn and is tuned to resonance. Where power gain is important at the lower frequencies, a different approach may be required to the antenna problem. For example, if several fixed antennas are used as the collectors for a direction-finder system, then in many applications a larger antenna structure is practical than in the case where the antenna is rotated. Thus at 50 Mc, a frequency at which a half-wave vertical element is about 10 ft high, several fixed vertical half-wave elements may be a feasible arrangement, whereas a rotating collector of even half the size may be impractical. A fixed antenna also has the advantage that rotating joints and drive mechanisms are avoided.

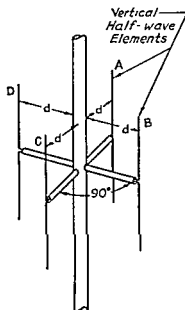


FIG. 10-46.—Fixed array of vertical elements around a mast.

A fixed array of vertical elements is particularly well adapted to installation around a mast, as illustrated in Fig. 10-46. In this illustra-

tion four vertical half-wave elements are mounted at a distance d from the mast and at 90-deg intervals. The directional characteristic of each antenna is substantially unidirectional since the mast acts as a reflector. By constructing a coupling system whereby each antenna is connected in turn to the receiver, a result is obtained that is equivalent to physically rotating one of the antennas around the mast. In other words, the directional characteristic is effectively rotated in azimuth. By observing the coupler orientation for maximum received energy, the direction or bearing of a source can be determined.

Such a fixed antenna array developed in the Radio Research Laboratory has a rotating coupler like that illustrated in Fig. 10-47. A trans-

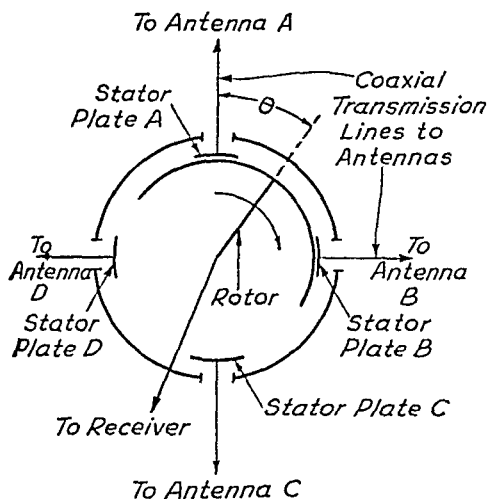


FIG. 10-47.—Cross section of coupler condenser for fixed antenna array.

mission line connects each of the four half-wave elements or antennas (A , B , C , and D of Fig. 10-46) to a stator plate of the coupler capacitor. A rotor plate is capacitance coupled to each of the stator plates in turn as it rotates.

A simple analysis of the system may serve to illustrate the principle of its operation. Assume that each of the antennas has a unidirectional pattern whose shape is a cosine function. Then energy from a distant source, in a direction at an angle ϕ with respect to a reference bearing, will be received by no more than two antennas at one time. The voltage developed across the terminals of antenna A is proportional to $\cos \phi$, while that in antenna B is proportional to $\cos (90^\circ - \phi) = \sin \phi$. By measuring θ , the angle of rotation of the rotor, as shown in Fig. 10-47, the resultant rotor voltage e can be expressed as

$$e = E_A \cos \phi \cos \theta + E_B \sin \phi \sin \theta \quad (10-26)$$

where E_A is the voltage on capacitor plate A when the source is at bearing

$\phi = 0^\circ$, and E_B is the voltage on capacitor plate B when the source is at bearing $\phi = 90^\circ$.

In Eq. (10-26) it is assumed that the spacing between the antennas is so small in terms of wavelengths that the difference in time of arrival of the incident wave at the two antennas can be neglected. In other words, the currents induced in the antennas by the passing wave are in phase. It is also assumed that the voltage on the rotor plate is a cosine function of the angle θ with respect to a stator plate.

If the system is properly designed the effective apertures of the antennas are the same, the transmission-line lengths are equal, and the shape of the coupler plates and their impedances to ground are identical. Then $E_A = E_B$ so that

$$e = E_A(\cos \phi \cos \theta + \sin \phi \sin \theta) \quad (10-27)$$

and

$$e = E_A \cos (\phi - \theta) \quad (10-28)$$

When the source direction or angle ϕ is constant, then the rotor voltage that is applied to the receiver is a function of $\cos \theta$ with a maximum when $\phi = \theta$. In a cathode-ray-tube presentation in which the radial deflection is proportional to the voltage e and the angular deflection equal

to θ , a cosine pattern will be traced on the tube face, as illustrated in Fig. 10-48. The maximum deflection is then at the angle ϕ , which is the direction or bearing of the source.

As indicated above, the coupler capacitor must be of such a design that the rotor voltage e due to a voltage E on the stator plate A can be expressed as

$$e = E \cos \theta \quad (10-29)$$

In a cylindrical capacitor with small difference in radii of rotor and stator, as in Fig. 10-47, the capacitance between the rotor and a stator is nearly proportional to the area common to

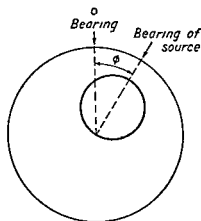


FIG. 10-48—Indicator pattern for fixed antenna array.

the two plates. The area in common must, in turn, be proportional to $\cos \theta$.

It is desirable that the impedance of the coupler capacitor be large compared with the antenna or receiver impedances to which it is connected, so that the phase of the voltage at the receiver will remain relatively constant with respect to that at the antenna as the coupler capacitor rotates. This condition does not correspond to that for maxi-

imum energy transfer from antenna to receiver, so that it is necessary to tolerate a loss in sensitivity.

The foregoing discussion concerns operation at a single frequency. The problem is more complex when operation over a band of frequencies is considered. For example, a limitation on the band width results from any change in directional pattern with frequency. A coupling capacitor can be designed for a particular antenna-pattern shape but will introduce deviations if the antenna pattern changes appreciably.

A fixed-antenna array for operation on horizontally polarized waves is sketched in Fig. 10-49, the principle of operation being the same as described above for the vertically polarized array.

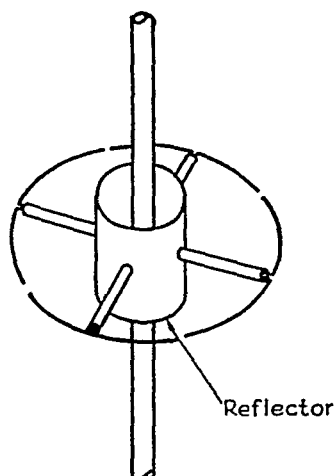


FIG. 10-49.—Fixed array of horizontal elements around a mast.

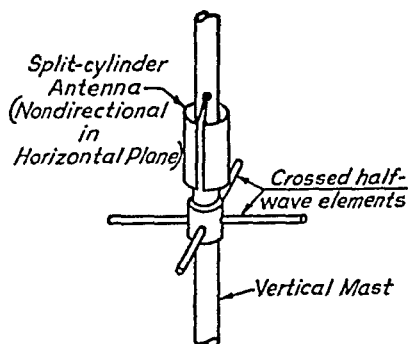


FIG. 10-50.—Fixed antenna array with cardioid pattern.

The four antennas of the arrays discussed above are assumed to be substantially unidirectional. An array of two bidirectional antennas mounted at right angles can be employed if the bidirectional pattern is not objectionable. The bidirectional characteristic can, however, be eliminated by coupling in the energy received from an antenna that is nondirectional in the horizontal plane. The arrangement shown in Fig. 10-50 illustrates an antenna of this kind for operation on horizontally polarized waves. The antenna consists of two crossed horizontal half-wave elements and a nondirectional split cylinder or loop antenna mounted around a vertical mast. By combining the outputs of the crossed elements and split cylinder properly with respect to phase and magnitude, in a coupler capacitor, a result can be obtained that is equivalent to physically rotating an antenna with a cardioid pattern.

10-42. Rotation-speed Considerations.—No discussion of rotating direction-finder antennas of the types described would be complete with-

out mention of the problem of intercept probability that arises when the source antenna is rotating; or when the receiver is being tuned; or, in the complicated general case, when the source antenna is rotating, the receiver is being tuned, the signal is pulsed, and its polarization is unknown. For the sake of illustration, attention will be directed to the case in which both the source and the spinner rotate and the frequency is fixed. An analysis of this problem reveals that there are *speed ratios* of the spinner to the source for which an intercept can be an impossibility (*e.g.*, a speed ratio of 1). The number of such ratios increases with decreasing beam widths of the spinner and source but is always finite. However, for speed ratios close to these values the interval of time necessary to ensure an intercept may be prohibitively long. In general, then, these speed ratios should be avoided. This can be accomplished with speed ratios such that the spinner rotates so much faster than the source that a bearing of the source is obtained at least once during each interval in which energy from the source is directed toward the spinner; or such that the spinner rotates so much more slowly than the source that a bearing of the source is obtained at least once during each interval in which the spinner is in position to receive energy from the source. The first condition is generally preferred, but for narrow beam widths may require very large speed ratios and hence high rotational speeds for the spinner; thus the mechanical problem of high-speed rotation may require consideration. Finally the fact that the source antenna rotates means that the field at the spinner fluctuates so that bearing errors as large as one-half of the beam width of the spinner may be encountered. These errors are minimized with speed ratios so large that several bearings are obtained each time the source is directed toward the spinner. The largest pattern on the indicator corresponds to the most nearly correct bearing.

10-43. Testing Techniques.—A number of considerations involved in the testing of the direction-finder antennas described are discussed in this section. To observe the directional characteristics of these antennas, a spacing of about 6 m between the spinner and the source antenna is satisfactory, provided certain precautions are taken. In some cases, to avoid incorrect results, it is necessary that the source antenna direct most of its energy toward the spinner. To illustrate this point, let us consider the test setup shown in Fig. 10-51a. Here the directional pattern of the source is such that most of the energy is directed toward the spinner under test. The fact that there are reflecting surfaces or walls W_1 and W_2 at either side of the spinner location does not materially affect the results, since the amount of energy radiated by the source in their direction is small compared with that radiated directly toward the spinner. Under these conditions the directional characteristic measured or observed for the spinner is as indicated by Fig. 10-51a.

Next, suppose that the directional characteristic of the source or transmitting antenna is not a single lobe, as shown in Fig. 10-51a, but is split into two lobes, as shown in Fig. 10-51b, with a minimum of energy radiated directly toward the spinner. The walls now play an important part since most of the energy from the source is directed toward them and then reflected toward the spinner. The measured or observed pattern of the spinner is no longer as shown in Fig. 10-51a, but now has the double-lobed appearance indicated in Fig. 10-51b, because in this case the energy received via reflection from the walls is large compared with that received directly from the source. Thus, to avoid this condition, it is apparent that the ratio of energy reaching the spinner directly to that reaching it via reflection should be large. To achieve this, the source should have a relatively sharp single-lobed pattern directed toward the spinner, and there should be no objects placed where they can reflect appreciable energy from the source to the spinner.

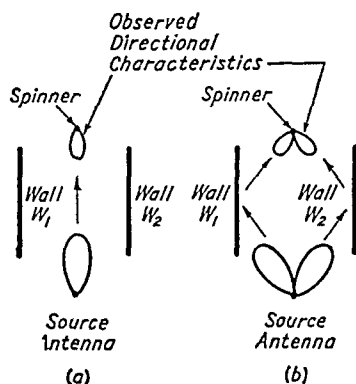


FIG. 10-51.—Effect of split source pattern and reflecting walls on observed patterns.

This example illustrates a rather obvious case. It is, however, sometimes not so obvious that a similar condition may exist in the absence of reflecting walls owing to the proximity of the ground or a reflecting plane below the equipment under test. If the directional characteristic of the source is relatively broad in the vertical plane so that approximately equal energies may reach the spinner position directly and via a ground reflection, the two waves may arrive at the spinner in opposite phase, making the field intensity at the spinner substantially zero. Under these conditions the field intensity has large variations or gradients with height. In other words, the vertical field distribution is nonuniform, or, to use an expression from optics, the "illumination" is not uniform. If we assume that the spinner is located at a position of substantially zero field, then the much larger field immediately below may cause mounting brackets or control cables beneath the spinner to reflect energy upward in amounts sufficient to produce observable effects. The field-intensity value, slightly to either side of this location or slightly nearer to or farther from the source, may not be the same. Thus, the field intensity may be nonuniform even over an area the size of the spinner. The spinner antenna will pass through these variations as it rotates, and as a result the observed spinner pattern may be considerably distorted.

The possible large variation in field intensity can be reduced materially in two ways: (1) by mounting the source and spinner at a distance above

the ground which is comparable to or greater than their distance from each other, and (2) by using a source with a directional characteristic such that very little energy is radiated toward the spinner via the ground reflection as compared with that radiated directly toward the spinner. The second method is readily applied at the higher frequencies where it is possible to construct a highly directional source, such as illustrated in Fig. 10-52a. At lower frequencies a highly directional antenna may be less practical. However, one may discriminate against the reflected wave by tilting the antenna to point upward, as shown in Fig. 10-52b. The upward tilt reduces the energy radiated directly toward the spinner somewhat, but may reduce the energy radiated toward the ground by a very much larger factor.

A test setup used at the Radio Research Laboratory employed a combination of a number of the above methods to provide relatively gradient-free fields at the spinner. The spinner and source were located

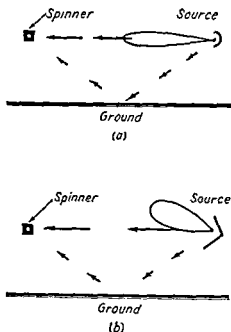


FIG. 10-52.—Relative location of spinner and source in test setup.

about 6 m above the roof of the building and about 6 m apart. The source antenna was a 120-deg corner-reflector antenna mounted to point upward (as in Fig. 10-52a) at an angle of 30-deg. The reflector aperture measured about 2.75 by 3 m. The radiating element could be adjusted to be about $\frac{1}{2}$ wavelength long at the operating frequency. The spacing from the apex of the corner reflector was also adjustable. By these adjustments a suitable directional characteristic of the antenna and reflector was maintained over the frequency range of 50 to 1000 Mc, producing relatively uniform illumination at the spinner location. The corner reflector could be rotated to either

a vertical or horizontal position so that waves of either polarization could be radiated. For frequencies above 1000 Mc the sources consisted of other types of antennas, such as paraboloidal reflector antennas and electromagnetic horns. These were mounted so that the maximum energy was radiated directly at the spinner and not above it.

If it is desired to test, for example, the disturbing effect of a vertical mast placed a few feet from a spinner, the mast and spinner can be

mounted so that the mast can be rotated around the spinner. In this way the effect of the mast on the spinner patterns can be observed as a function of the angle with respect to the source direction. For significant results in this type of test, it is necessary that the field intensity be relatively gradient free over the entire region between the mast and the spinner and even beyond. If the mast is in a strong field and the spinner in a weak one, energy reflected from the mast may predominate and produce incorrect results. To determine if the region is gradient free, a half-wave element can be used to explore the region while the received field intensity is observed with a receiver.

In comparing the power gain of a spinner antenna with a resonant half-wave element, the element can be substituted for the spinner at the spinner location and the power output from the r-f generator to the transmitting antenna varied until the receiver indicates the same power input as before. The conditions for the reference element measurements should approximate free space. The general arrangement is shown in Fig. 10-53. The ratio of the two transmitter powers is then a measure of the power gain. Because of the fact that the receiver may not present a reflectionless termination to the r-f transmission from the antenna, the measured power gain may be found to be a function of the receiver or of the transmission-

line length when there is an impedance mismatch at the antenna to transmission-line junction (see last part of Sec. 10-16). To reduce the fluctuations in measured power-gain values as a function of frequency, a device may be inserted in the r-f transmission line so that the length can be adjusted at will. Such a device is commonly called a *line stretcher*. By adjustments of the line stretcher, maximum and minimum power gains can be measured for a given frequency. For many purposes, it is convenient to take the average of these two values, giving an *average power gain*. For example, the effective apertures presented for the corner spinner in Fig. 10-28 were derived from measured average power gains. Any variation in attenuation due to changes in transmission-line length produced by a line stretcher can usually be neglected if the total length of transmission line is large compared with the change in length.

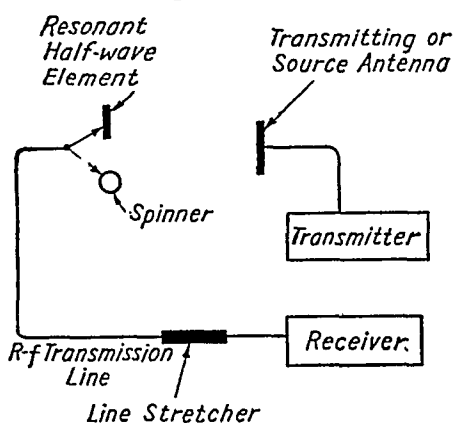


FIG. 10-53.—Test setup for comparing the power gain of a spinner-antenna with a resonant half-wave element.

CHAPTER 11

INDICATORS FOR DIRECTION FINDERS

By W. McGUIGAN

11-1. The Problem.—In Chaps. 9 and 10 it has been pointed out that at frequencies above 50 Mc direction-finder systems using relatively small rotatable directional antennas are possible. In fact, as the frequency becomes higher, the advantages of such antennas become greater. The problem then remains of providing the direction-finder operator with sufficient equipment for determining the direction of signal arrival. The portion of the direction-finder system not included in either the antenna or receiver that provides such data to the operator is called the *indicator*.

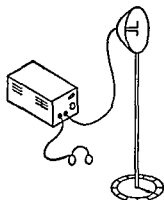


FIG. 11-1.—Simple direction finder. Direct indication

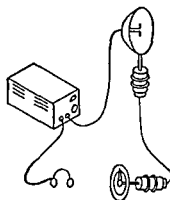


FIG. 11-2.—Simple direction finder. Remote indication

In an elementary form, a microwave direction finder may be comprised of a rotatable antenna having a suitable dial and an associated receiver, as shown in Fig. 11-1. The antenna is adjusted to give maximum response, and the bearing is read on the dial attached to the antenna shaft. Although simple, this system is not particularly convenient. For instance, it might be desirable to place the indicator at a point remote from the antenna.

The function of an indicator unit can be resolved into two major parts: (1) that of presenting information as to the angular position of the antenna and (2) that of presenting information as to the strength of the received signal. When both types of information are properly presented in combination, the bearing of the source may be deduced. We shall

first consider some of the methods of transmitting information as to the angular position of the antenna.

11-2. Remote Indication.—An elementary direction finder incorporating a servo system that permits remote control of the antenna is shown in Fig. 11-2. As in the previous system, when the antenna has been adjusted for maximum response, the bearing is indicated by a pointer that rotates in synchronism with the antenna. By turning the dial at the operating position, the operator excites the synchro (or selsyn) system shown, causing the antenna to rotate by a similar amount. This system is simple in that the same synchros that rotate the antenna also repeat the bearing information.

11-3. Need for Automatic Direction Finding.—Neither of the systems described above is suited to more than limited applications. The operation is dependent upon the operator's ability to hear the signal; signals with supersonic or subaudible modulation cannot be located, and the over-all sensitivity of the system is a function of the operator since the a-f response varies among individuals. Furthermore, at frequencies above 50 Mc one commonly encounters signals from sweeping sources, such as radars. Because these signals are not heard except for brief intervals, the operator may be unable to align the antenna in the proper direction in the brief time available. Where the signal must be located with the greatest possible accuracy in the minimum possible time, the operations performed by hand in Fig. 11-2 must be made automatic.

Before discussing the construction of an automatic direction finder, we shall consider briefly some of the operations involved in the determination of a bearing. For example, the receiver must first be tuned to the proper frequency. If the bearing is obtained by listening, the antenna must then be rotated back and forth until the response is a maximum (or a minimum in the case of a null-type system). To be sure that the antenna is properly aligned, it is necessary to compare the signal strengths on either side of this central setting. Thus, in taking a bearing, the correct alignment may be determined not so much from the position of maximum response, as from two positions of equal response that are presumed to be equally spaced from the center. To simulate the hand operations, an automatic system must therefore rotate or sweep through the direction of the source.

As mentioned in Chap. 10, if the direction finder rotates automatically, it is desirable that the speed be considerably greater than that of the sweeping source. This may require speeds of the order of a few hundred rpm for the rotating antenna, and under these conditions it is apparent that some form of automatic presentation is required.

11-4. Rotating-disk Indicator.—For automatic presentation, the system of Fig. 11-2 must be modified. The simplest manner of combin-

ing the signal with the directional data is shown in Fig. 11-3. This is particularly suitable where frequencies are high enough to permit a

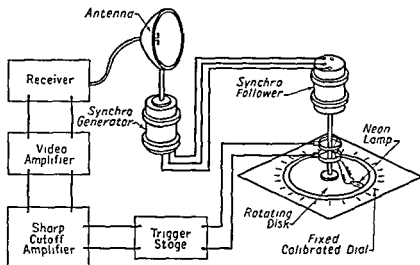


FIG. 11-3.—Rotating-disk indicator. The trigger stage permits uniform voltage to be supplied to the neon lamp during the signal interval. The sharp cutoff amplifier permits a sharpening of the lighted sector by a manually adjustable gain control.

highly directive antenna. As the antenna rotates, a signal will be heard in the receiver when the antenna is pointed at the signal source. By coupling a small bulb to the output of the receiver, a light will be seen whenever the signal is received. The flashing bulb is then mounted on a disk or dial which rotates in synchronism with the antenna.

Since the signal is received only when the antenna is pointed in the direction of the signal, only a certain sector of the rotating disk will appear visible. The dial appears as shown in Fig. 11-4. As illustrated the bulb attached to the rotating disk showed five flashes as the antenna swept past the direction of a pulsing source or radar. In general, the illuminated sector of the rotating dial is proportional to the beam width of the direction-finder antenna. The spacing of the individual flashes is the angle through which the direction-finder antenna turns between successive pulses of the received radar. If the modulation

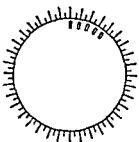


FIG. 11-4.—Presentation of rotating-disk indicator. Flashes caused by successive pulses in the received signal appear simultaneously due to rapid rotation of the disk and persistence of vision. A bearing of 20 deg is indicated.

of the signal were steady, the lamp would glow throughout the sector.

With this type of indicator it is desirable that the operator be able to use earphones or other means of identifying the signal before taking a

bearing, because the flashing bulb does not in itself permit easy identification of the signal. Furthermore, at relatively low frequencies, where the problem of designing a highly directional antenna sufficiently small to be rotated at two or three hundred rpm is impractical, the broad lobe from an antenna of practicable dimensions would cause the lamp to glow over a sector too great to be accurately bisected. This indicator, therefore, is suited only to frequencies at which the antenna beam width is of the order of 15 deg or less.

11-5. Pulse Indicators. Magnetic-coil Method.—For a direction finder that must operate at lower frequencies and utilize the information from a less directive antenna array, another type of indicator is more desirable. An indicator with a bearing presentation similar to that of Fig. 11-8 (although not necessarily with broadened tips) has shown considerable merit. Such a presentation can be produced by a system similar to that shown in Fig. 11-5, except that the rectifier-filter stage is omitted and the beam modulation comes directly from the video amplifier. The synchros are used to drive the deflection coils of a cathode-ray tube, which therefore rotate in synchronism with the antenna. The output of the receiver is amplified and applied to the rotating coils. The modulation of the signal then causes radial deflections on the screen of the cathode-ray tube. As the antenna rotates, the signal in the receiver varies in amplitude, becoming strongest when the antenna is pointed directly at the source. Since the direction in which the trace moves from the center of the screen depends upon the instantaneous direction of the direction-finder antenna, the pattern of Fig. 11-8 represents a bearing of 20 deg (as determined by the bisector of the pattern).

11-6. Envelope Tracers. Magnetic Scan.—A disadvantage of the magnetic-coil method indicator is the difficulty of designing deflection coils that will function efficiently, not only on communications and voice-modulated signals, but also on radar signals having pulses as narrow as $1\mu\text{sec}$ or less. It is not always possible to design, for limited space, deflection coils with suitable band-pass characteristics and sufficient field strength to give adequate deflection. Effective presentation of the antenna response can be obtained by amplifying and rectifying the receiver output and connecting the rectified output to the rotating deflection coils of the cathode-ray tube. The rectified deflection current that is proportional to the signal strength traces an envelope of the antenna pattern as the antenna rotates. This system is shown in Fig. 11-5. The pattern on the indicator screen is shown in Fig. 11-6. For identical antennas and circuits of comparable linearity, this is merely the envelope of the pattern of Fig. 11-8.

If the envelope traced on the screen is to be smooth and relatively independent of the modulation frequency of the incoming signal, the audio

or video signal must be not only rectified but filtered. The rectifier output may be dependent on the type of modulation. A voice-modulated

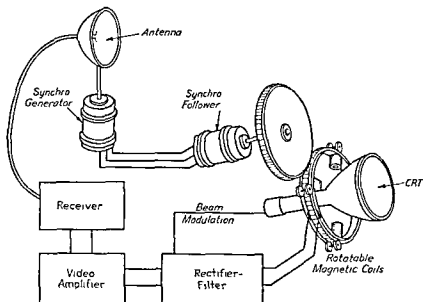


FIG. 11-5 — Envelope-tracing indicator

signal may produce a considerably greater deflection voltage than would a few relatively narrow pulses from a radar station of the same signal intensity. Thus the sensitivity of the direction-finder system may

vary considerably with the type of signal. The rectifier may also have a time constant that subjects the system to variations in accuracy if the rotation speed of the direction-finder antenna is changed. Finally, the time constants of the filter may prevent the rectified current from being exactly proportional to the instantaneous signal strength. The result may be an asymmetrical pattern, as shown in Fig. 11-7b. Figure 11-7a shows for comparison an undistorted pattern. Therefore when filters are used in the indicator unit and the antenna system of the direction

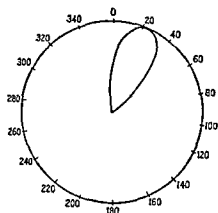


FIG. 11-6 — Presentation of envelope-tracing indicator. Bearing approximately 22 deg

finder is run at various speeds, there may be an apparent shift of the bearing. Thus filters should be used only with scanning circuits having frequencies below the cutoff frequency of the filter.

A system of the envelope-tracing type has limitations in that the deflecting current that results if two interfering signals are heard causes an indication of a bearing somewhere between the bearings of the two signals and gives the correct bearing of neither. Naturally, if the sources have widely different bearings and the antenna of the direction finder has sufficient directivity, the resolving power of the antenna will enable the signals to be presented separately, and no interference will result. But the indication will be incorrect for two signals of comparable magnitude and nearly the same bearing. If it is possible for the operator to use headphones or other means of signal analysis, the interference of such comparable signals can be detected, and the objection is not serious.

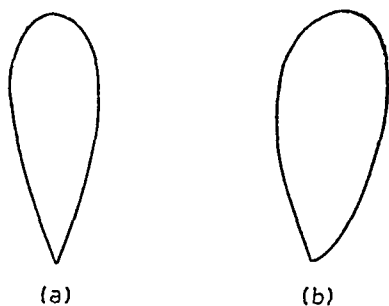


FIG. 11-7.—Distortion of pattern by indicator circuits: (a) normal pattern; (b) distorted pattern due to long-time constants in filter of envelope-tracing indicator.

11-7. Electrostatic Pulse Indicator. Capacitor Scan.—A widely used type of cathode-ray-tube presentation, in which the envelope is similar to that of the system described in Sec. 11-6 but which permits independent bearings on stations with nearly the same bearing and the

same frequency, is shown in Fig. 11-8. This is the presentation of a radar signal, and the individual spokes represent successive pulses from the radar. As in the previous cases, the bearing is determined by bisecting the lobe and reading the angle on a cursor around the edge of the screen. One system for presenting the pulses in this manner is shown schematically in Fig. 11-9. Here the rotating antenna is connected to a receiver, the output of which is connected to the indicator unit. A

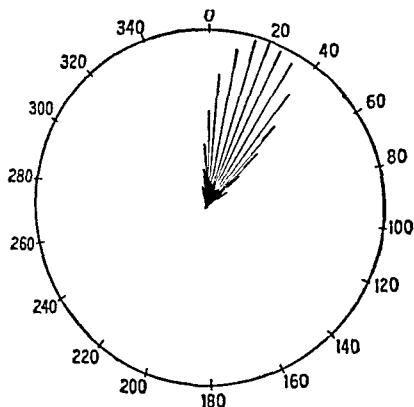


FIG. 11-8.—Pulse presentation of electrostatic indicator.

scanning circuit within the indicator unit is coupled to the antenna by synchros so as to rotate in synchronism with it. As the antenna rotates and the received signal intensity varies, the indicator sweep circuit also rotates and successive pulses received by the system vary in amplitude, as shown. The pulses of maximum length represent the point at which the antenna is directed toward the source. The essential parts of this system are described in following sections.

11-8. Scanning Circuit.—The signal from the receiver is coupled to an amplifier with band-pass characteristics designed for the type of signal it is desired to receive. The output of the amplifier is applied to the rotor of a

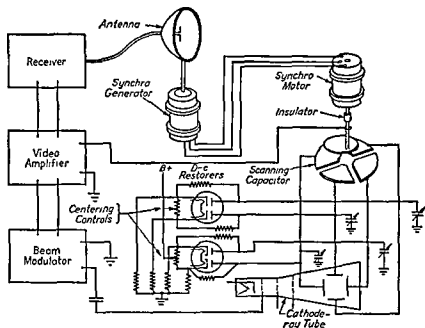


FIG. 11-9.—Schematic diagram of scanning circuit for electrostatic indicator.

scanning capacitor with four independent stators. Details of this capacitor may be seen in Fig. 11-10. These stators, with the rotor, form the first capacitors in four parallel capacitor-divider networks to ground

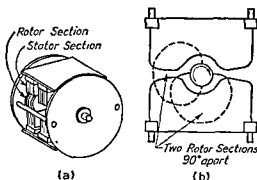


FIG. 11-10.—Scanning capacitor for electrostatic indicator: (a) isometric view; (b) end view showing 90-deg relation of rotor plates.

(see Fig. 11-9). As the rotor turns, and the coupling is varied to each of the four networks, the difference voltage between opposite pairs of these dividers varies in a sine-cosine relation by virtue of the design of the capacitor plates. The difference voltages are then applied to the two

pairs of plates in the cathode-ray tube, producing a deflection the amplitude of which depends on the intensity of the received signal, and the direction of which from the center of the screen depends on the angular position of the rotor in the scanning capacitor. Because the capacitor rotates in synchronism with the antenna, the direction of the instantaneous trace corresponds to the instantaneous position of the antenna.

11-9. Restorer Circuit.—Because for modulated signals the receiver output may swing alternately positive and negative with respect to ground, it is necessary to put in a restoring device to shift the reference voltage of the deflection circuit and thereby prevent the traces on the cathode-ray tube from passing through the center of the screen. This is accomplished by the two double-diode restorer tubes shown in the scanning-circuit diagram. The restorer circuit not only causes the traces to originate at the center of the screen, but also facilitates the injection of horizontal and vertical beam-centering voltages for the cathode-ray tube. From the schematic diagram of Fig. 11-9 it can be seen that when negative voltages are applied to the rotor of the capacitor the diodes constitute a high impedance and have no effect on the circuit. With a sine wave, or other signal applied to the rotor, the positive components of the signal will cause the diodes to conduct, charging the small trimmer capacitors attached to each of the CRT deflecting plates. Since the resistors shunting the diodes are large, usually at least 4 megohms, the trimmers will retain their charge after the rotor signal swings negative. Thus the effective negative deflection voltage is measured with respect to the most positive peak of the signal, permitting twice the deflection voltage that would be possible if the diodes were used as "clippers."

11-10. Beam Modulator.—An auxiliary stage between the video amplifier and the cathode-ray tube is shown in Fig. 11-9. The purpose is evident if we recall from Fig. 11-8 that the deflections represent only a small portion of the duty cycle. The remainder of the time the spot on the cathode-ray-tube screen would be at the center and the screen would tend to become burned. Further, the brilliance of the spot at the center would interfere with the visibility of the actual pattern, particularly when the pulses are short and the traces are faint. The auxiliary stage converts the negative signal from the video amplifier into a positive voltage that can be applied to the brilliance grid of the cathode-ray tube at the same time that the deflection voltage appears in the sweep circuit. This *beam modulator* permits the center of the screen to be blanked out during the intervals between pulses.

11-11. Amplifier Circuits.—An indicator of this type presents different types of pattern for different types of pulsed signals, as shown in Fig. 11-11. The repetition rate determines the number of pulses seen during one revolution of the direction-finder antenna. This type of

presentation facilitates a limited amount of pattern analysis and also the determination of the bearing of a second signal of the same frequency as the first. Both signals can be seen separately, even though they may have the same or nearly the same direction, as shown in Fig. 11-12.

The design of the amplifier may have considerable influence on the presentation in this type of indicator. The patterns shown in Figs.

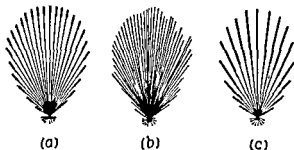


FIG. 11-11.—Appearance of various signals on electrostatic indicator (a) short pulses, high p.r.f; (b) speech modulation; (c) long pulses, low p.r.f

11-8, 11-11, and 11-12 are the result of an indicator with an amplifier that distorts the pulses, as shown in Fig. 11-13. This makes the tips of the pulses appear as lines, the length of which depend on the pulse length. An experienced operator can then tell the approximate pulse length and, from the spacing of the pulses, the approximate repetition rate. This facilitates identification of the source.



FIG. 11-12.—Two signals on electrostatic indicator.

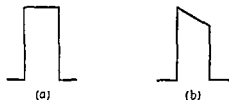


FIG. 11-13 —Pulse distortion in electrostatic indicator: (a) input signal; (b) output signal, showing sloping top of pulse.

Under some circumstances it may be desirable to make the pulses perfectly square before applying them to the deflection circuits. But prudent adjustment of the beam modulator and proper pulse lengths can yield a pattern similar to that shown in Fig. 11-14. Again, the tops of the pulses are made to slant slightly, but the main part of the radial trace disappears. This leaves the central portion of the pattern open, which is highly desirable when there is a possibility of receiving several signals on

the same frequency. Simultaneous bearings on two stations having nearly the same direction then appear as shown in Fig. 11-15.

One difficulty with a presentation of this type is that successive pulses of the received signal are spread over the screen of the cathode-ray tube.



FIG. 11-14.—Pulse presentation. Critical adjustment of beam modulator.



FIG. 11-15.—Presentation of simultaneous bearings. Two signals more visible with modulator adjustment.

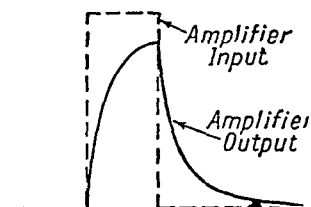


FIG. 11-16.—Undesirable distortion in electrostatic indicator.

In addition to not falling on top of each other, the individual pulses are stretched along a radial line. For this reason the extreme brilliance that would result if successive pulses appeared on top of each other is not possible. In practice one finds that certain types of pulses must be distorted to increase the pattern brilliance. Another way of stating the problem is to say that the amplifier most desirable for certain types of signals would distort rather than faithfully reproduce the input voltage from a receiver having little or no distortion. If the pulse is reshaped as indicated in Fig. 11-16, it is evident that the maximum brilliance will appear along the portion of the pattern closest to the center of the screen. Because with such a pattern the operator depends on the envelope for interpretation, the brilliance obtained by pulse distortion will not be in the portion of the trace that aids in the determination of the bearing. Further, the additional brilliance at the center may obscure the pattern of a second signal, as shown in Fig. 11-17.

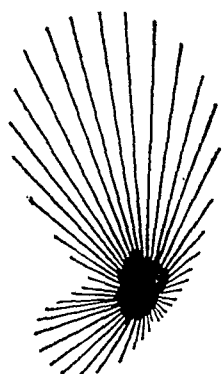


FIG. 11-17.—Signals blurred by electrostatic indicator having distortion of Fig. 11-16.

11-12. Pulse Stretching.—With certain types of radar signals the received pulses are of sufficiently short duration to make their trace on the indicator screen indistinct. This is particularly true for pulse lengths of the order of $1\ \mu\text{sec}$ or less. Pulses that have lengths of the order of $5\ \mu\text{sec}$ or more can be made more visible by the proper design of an amplifier, as pointed out in the previous section. But for shorter pulses a new circuit is desirable that produces longer square pulses whenever it is triggered by the received signal. Such a circuit may be called a *pulse*

stretcher and has no function other than that of increasing the screen intensity. A typical arrangement is illustrated in Fig. 11-18. The one property that distinguishes this circuit from conventional stretcher cir-

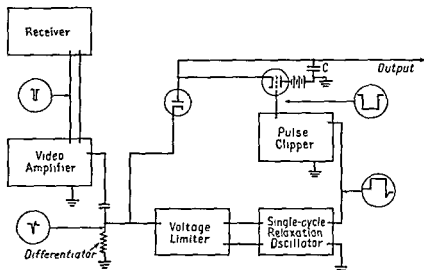


FIG. 11-18.—Block diagram of pulse stretcher for electrostatic indicator

cuits is that the output pulse must have an amplitude proportional to the amplitude of the input pulse.

Details of this circuit may be seen in Fig. 11-19. The pulse from the receiver is amplified and applied to a clamping circuit that consists of a

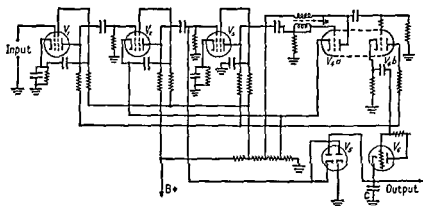


FIG. 11-19.—Schematic diagram of pulse stretcher. V_1 and V_2 , video amplifiers; V_3 , voltage limiter; $V_4 a$, relaxation oscillator; $V_4 b$, pulse clipper; V_5 , clamping diode; V_6 , gate triode.

diode in series with the capacitor C . A resistance-capacitance arrangement at the output of the amplifier differentiates the pulse, so that the following circuits are triggered by the leading edge of the pulse and the duration of the output pulse is independent of the input pulse length.

The same voltage that is applied to the clamping circuit and charges *C* is also sent through a voltage limiter to a pulse circuit. The output of the pulse circuit is coupled to an inverter stage that also serves to improve the waveform of the pulse. The reshaped pulse is then applied to the grid of a triode in such a manner that when the negative output of the inverter is removed the triode conducts, discharging the capacitor *C*. The output voltage is measured across *C*, which is charged to an initial amplitude by the incoming signal. This voltage remains until the end of the pulse from the pulse circuit, when the capacitor is discharged.

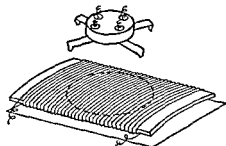
A general-purpose radar direction finder might encounter signals with pulse lengths varying from 0.2 μ sec or less to 80 μ sec and pulse-repetition rates between 25 and 4,000 cps. As the pulse rate increases, the spokes due to successive pulses (see Fig. 11-8) on the screen of the indicator come closer together, until they form a solid pattern. This is undesirable if it is necessary to observe at the same time two signals having nearly the same bearing. Also, the uniform output of the pulse stretcher tends to hide the characteristics of the individual radar, making identification more difficult. In the presence of a high noise level the stretcher will be triggered by the noise rather than the signal, and the signal will be completely obscured. For these reasons, if a pulse stretcher is included in a general-purpose radar direction finder, it should be possible to switch it out when normal operation is desired.

11-13. Cathode-ray-tube Screens.—The type of cathode-ray-tube screen to be used with an indicator of this type should be selected to aid the pattern analysis. If one assumes that a persistent screen is desirable so that several successive sweeps of the direction-finder antenna can be used to build up a pattern of considerable intensity, then a P7 or similar screen of great persistence may be employed. A difficulty with this screen, or with most two-color screens, is that the single trace in one-pulse presentation does not sufficiently excite the primary coating to excite the persistent materials of the screen. The trace on a P7 screen appears rather as a bluish sparking than as the yellow glow. This is particularly true for pulses of the order of 1 μ sec or less. Another consideration is that excessive persistence may cause the pulses from so many successive antenna sweeps to pile up on top of each other as to confuse the operator, particularly if the signal is varying in amplitude or if more than one signal is present. Experience has shown, therefore, that a screen of the P2, or medium-persistent variety, is more desirable. Such a screen can be excited by a relatively narrow pulse, and at the same time the persistence is sufficient to permit only one sweep of the direction-finder antenna to be presented at a time. With regard to cathode-ray-tube screens a compromise is often made for inexpensive lightweight indicators, because screens having considerable persistence, *i.e.*, either

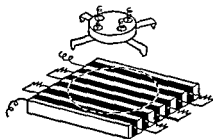
the P2 or P7 variety, require high accelerating voltages. Since these may require twice the voltage for a simple nonpersistent screen, the extra space needed for transformers and high-voltage insulation is often eliminated at the expense of a partial loss in signal analysis. Bearings can be taken with equal accuracy under similar lighting conditions with any of these screens.

11-14. Electrostatic Indicator. Sine-potentiometer Scan.—In many instances the direction finder must function on signals having modulation other than the pulses peculiar to radar stations. For example, there are

communications signals, beacons, and various navigation devices that may have speech, tone, or supersonic modulation. Referring to Fig. 11-9, we observe that the scanning circuit consists of a quadrant capacitor that forms a variable capacitance-divider network with each of four trimmer capacitors. Since each of the trimmers has a capacitance of the order of $50 \mu\mu\text{f}$, it is evident that this scanning circuit will have a very high impedance with low-frequency audio modulation. Furthermore, the return path for the restorer currents must consist of resistors having impedances that are high compared with the trimmers which they shunt. In practice these resistors are seldom less than 4 megohms, tending to a



(a)



(b)

FIG. 11-20.—Sine potentiometers: (a) wire-wound type; (b) sandwich type.

very high-resistance leakage path between the deflecting plates of the cathode-ray tube and ground. At low audio frequencies the impedances of the trimmer capacitors become comparable with that of the diode-return path. This adversely affects the accuracy of the indicator because of differences in the time constants in the various quadrants.

These difficulties can be eliminated by substituting a sine potentiometer for the quadrant capacitor and trimmer arrangement of Fig. 11-9. Several designs of sine potentiometers have been proposed by the Radiation Laboratory. Two of these are illustrated in Fig. 11-20. One consists of a coil of resistive wire wound on a plastic card. The windings are formed so as to bend the card slightly, keeping the coil under tension at all times. Four pickup brushes on a rotor then travel on a circular path

over the coil, producing sine and cosine voltages between opposite pairs of brushes. The brushes are connected to the deflecting plates of the cathode-ray tube of the indicator unit.

The second type of sine potentiometer illustrated is a sandwich fabricated from alternate layers of beryllium-copper and cellophane of nearly equal thickness. The final form is machined to form a smooth surface against which the four contacts may ride. Because of the relatively low resistance of the conducting layers, a strip of resistive material is used along the edge to give the desired resistance to the potentiometer. To reduce brush noise the conventional practice is to use a small resistance-capacitance filter in series with each of the four contacts. A typical arrangement is shown in Fig. 11-21. The filters are not adaptable to video signals, and for this reason the sine potentiometers are not used with direction finders that operate on radar signals.

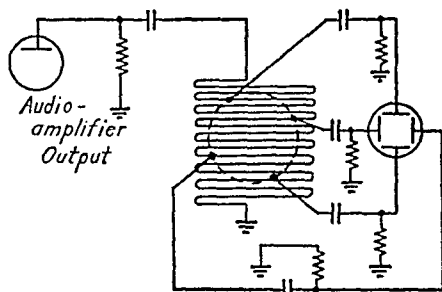


FIG. 11-21.—Schematic diagram of sine-potentiometer scanning circuit for an electrostatic indicator.

11-15. Electrostatic Indicator. Electronic Scan.—In the pulsed envelope indicators thus far mentioned, the relative efficiency of the scanning circuit is an important consideration. Whether a scanning capacitor or a sine potentiometer is used, the fact that a difference voltage is obtained between two voltages of equal sign but of different magnitudes causes the resultant deflecting voltage applied to the cathode-ray tube to be seldom more than 30 per cent of the output voltage of the indicator amplifier. For this reason, indicators of the type shown in Fig. 11-9 are likely to have amplifiers and associated power supplies that are relatively large. Another consideration is that a scanning circuit containing mechanically moving parts may not be capable of following the antenna rotation accurately, either at varying speeds or when sector scan is desired. A considerable saving in space and weight can be effected in addition to improved operation by obtaining the sine-cosine voltages directly from the antenna system. These voltages can then be used to form a variable bias that controls the deflections of an electronic scanning circuit, as shown in Fig. 11-22. The signal voltage is fed into an amplifier rather than a voltage-loss device, resulting in considerably greater efficiency. Also, the scanning voltage from the antenna synchro is used only as a bias, so relatively low voltages are encountered in the sweep circuit.

The sweep circuit is shown schematically in Fig. 11-22. A two-

phase synchro, mechanically coupled to the rotating antenna, has its rotor excited by a voltage from an oscillator in the indicator unit. The sine-cosine outputs of the synchro are then fed back to the indicator unit

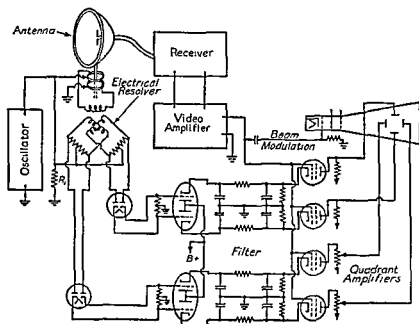


FIG. 11-22.—Electrostatic indicator with electronic scan. The sine and cosine outputs of the synchro are rectified, filtered, and applied as variable cathode bias to the four quadrant amplifiers. The CRT trace is deflected toward the quadrant or quadrants having the most positive cathode bias.

and applied to a discriminator circuit. The operation of this circuit may be more clearly seen in Fig. 11-23. The voltage from the synchro is fed to two diodes in push-pull through the transformer T_1 . At the same time another voltage, from the same source as the synchro voltage,

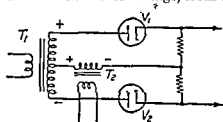


FIG. 11-23.—Phase-amplitude discriminator.

is applied to the diodes as a reference voltage through the transformer T_2 . For instantaneous voltages with the polarity shown the resultant voltage applied to V_1 is larger than that applied to V_2 , giving a difference voltage across the ends of the load resistors which can be coupled to the following circuit. As the amplitude of the synchro voltage changes, the difference voltage at the output of the diodes also changes. When the antenna synchro is turned 180 deg, the voltage at the output of T_1 reverses its phase. The voltage resulting from the sum of the reference voltage and the output of T_1 then yields greater voltage across the lower diode V_2 , reversing

the following circuit. As the amplitude of the synchro voltage changes, the difference voltage at the output of the diodes also changes. When the antenna synchro is turned 180 deg, the voltage at the output of T_1 reverses its phase. The voltage resulting from the sum of the reference voltage and the output of T_1 then yields greater voltage across the lower diode V_2 , reversing

the polarity of the voltage across the ends of the load resistors. The voltage at the output of this circuit is therefore proportional to the amplitude of the voltage from the antenna synchro and has a polarity dependent upon the phase of the antenna synchro voltage.

In Fig. 11-22 the output of the diode discriminator circuit is coupled to a cathode-follower stage in order to match the comparatively high impedance of the diode outputs to the cathode impedances of the quadrant amplifiers. The output of the cathode-follower circuit contains a resistance-capacitance filter. To reduce the time constants of the filter, the frequency of the oscillator is chosen as a compromise representing maximum efficiency of the synchro and minimum loss in the antenna cable. Thus, at a frequency of a few hundred cycles per second, a filter can be constructed that does not tend to eliminate the ripple corresponding to the frequency of rotation of the antenna. This is important in that a filter which attenuates this ripple will distort the envelope of the cathode-ray-tube pattern, as shown above in Fig. 11-7b.

In this indicator, as in the previous one, a beam-modulator tube is used for the purpose of applying a small portion of the signal voltage to the brilliance grid of the cathode-ray tube, to increase the brilliance of the trace during periods of screen deflection, to blank out the center of the screen, and to increase the contrast.

11-16. Radial Presentation.—In the cathode-ray-tube indicators thus far discussed, the method of presentation is to indicate a directional signal by deflections from the center of the screen. The methods of scanning are both mechanical, using rotating magnetic deflection coils, scanning capacitors, or sine potentiometers, and electrical, using a variable bias arrangement to control the direction of the trace. Because the deflections are always between the center and the rim of the CRT screen, it would appear that a more obvious approach to the indicator problem might be to use a radial-type cathode-ray tube, with a deflecting anode along the axis of the tube. In Fig. 11-24, for example, we see a directional presentation in which the sweep voltage causes the trace to travel around the outer edge of the screen as the antenna revolves. The instantaneous angular position of this spot thus corresponds to the direction in which the antenna is pointed. As the antenna rotates and points in the direction of the signal source, the output of the receiver can be applied to an anode located along the axis of the cathode-ray tube.

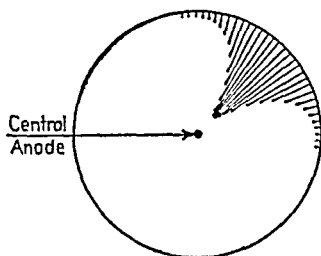


FIG. 11-24.—Radial presentation. Positive deflection voltages on the central anode pull the sweep trace from the rim of the screen during pulses. The pattern is bisected to determine the bearing.

presentation cannot be inverted to give a presentation similar to that shown in Fig. 11-8 by using a scanning trace that travels in a small circle near the center of the screen and negative voltages on the deflecting anode, because (1) the angular error due to distortion of the electron beam by the central anode is too great for a direction finder, and (2) the fact that the voltage gradient around the radial-deflecting anode is not linear tends to reduce the apparent directivity of the antenna.

11-17. Electrostatic Indicator. Synchro Scan.—A better approach to the problem of a compact lightweight indicator is shown schematically

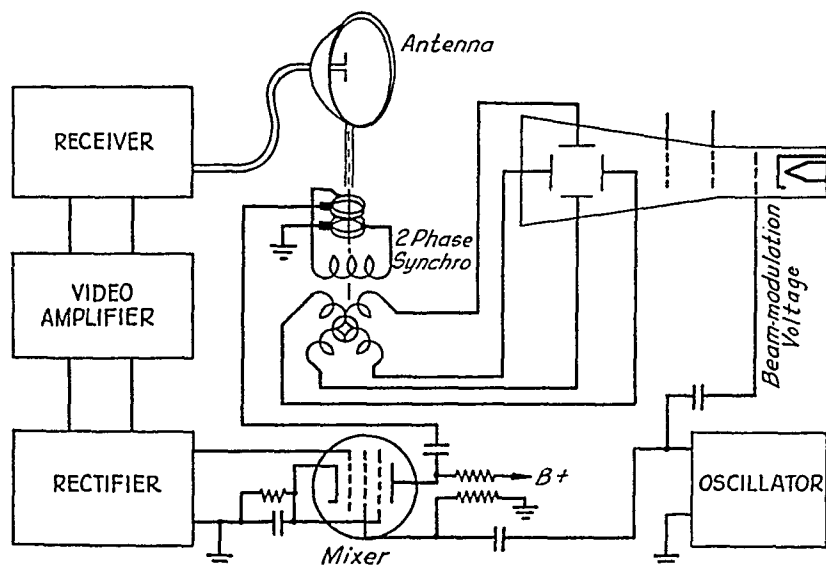


FIG. 11-26.—Electrostatic indicator with synchro scan. The signal from the receiver is rectified and used as a bias to control the amplitude of the voltage in the rotor of the antenna synchro. The CRT trace then has a direction corresponding to the angular position of the antenna and a length dependent on the signal strength. The antenna and its synchro turn in a 1:1 ratio.

in Fig. 11-26. The principle of operation of this indicator is to apply the output of the two-phase synchro directly to the plates of the cathode-ray tube. The rotor of the synchro is excited by an oscillator-amplifier with output amplitude proportional to the strength of the received signal. This proportionality is brought about by rectifying the receiver output to produce a bias voltage that is applied to the first grid of a mixer tube. The signal grid of the mixer is driven by a local oscillator. The output of the mixer, which is applied to the rotor of the antenna synchro, then consists of the oscillator voltage modulated by the signal strength at the output of the receiver. Since the voltages induced in the two secondaries of the antenna synchro are in phase, the trace described by these voltages on the cathode-ray-tube screen is a line passing through the center. To

indicate the direction in which the antenna points, the back half of the trace must be suppressed by connecting the oscillator output so as to modulate the brilliance grid of the cathode-ray tube. Thus only the half of the trace representing the positive swing is visible. The screen pattern, shown in Fig. 11-27, consists of 2500-cps sweeps modulated by the signal strength as the antenna turns. The 2500-cps value is arbitrary and can be varied widely.

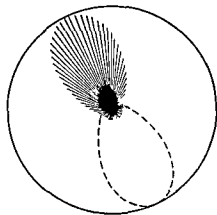


FIG. 11-27.—Presentation of synchroscan indicator. Individual deflections consist of sinusoidal half waves at frequency of local oscillator. Dotted line indicates half of traces eliminated by beam modulation.

11-18. Double-line Indicator.

All the indicator units described thus far were designed for a direction finder with an antenna that rotates and has maximum response when pointed at the source of the signal. In a direction finder with a null-type array, however, the bearing is determined when the antenna is adjusted to give a minimum response. Such systems are used at lower frequencies where

the indicator may consist of little more than a calibrated dial that can be read after the antenna has been adjusted manually. Above 50 Mc, however, signals with inaudible modulation are sometimes encountered and visual presentation is desirable.

With a null-type direction finder the null is most accurately located by rotating the antenna array back and forth and comparing signal strengths on the two sides of the null. In simple systems this operation is manual, and the signals are compared aurally. When greater accuracy is required or the modulation of the signal is unsuitable, provision may be made for a visual indicator and the antenna array made to sweep back and forth automatically. The data can then be presented on the face of a cathode-ray tube in the manner shown in Fig. 11-28. The two lines represent the signal strengths on the two sides of the null. If the antenna system is displaced to one side, one of the lines increases its length. When the array is aligned to sweep an equal angle on each side of the null, the lines are of equal length. This presentation can be accomplished by means of a relay located on the antenna-drive unit that closes

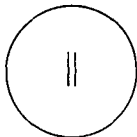


FIG. 11-28.—Presentation of two-line indicator. The direction of the signal is obtained when the two lines are of equal length. When the antenna system is displaced one of the lines becomes longer.

on one side of the center of the sweep and stays open on the other. When the relay closes, the horizontal positioning voltage for the cathode-ray tube is changed, so that the vertical trace representing the amplitude of the signal at any instant appears on one side or the other, depending on which half of its sector the antenna is sweeping.

11-19. Comparison of Presentations.—The null-type direction finder described in Sec. 11-18 is at a disadvantage when used on intermittent or sweeping signals, which are observed for only 3 or 4 sec at a time, perhaps twice a minute. Such systems can therefore be disregarded except for compact lightweight direction finders used for special purposes. For automatic direction finders, however, several methods of either the maximum or null type of cathode-ray-tube indication might be considered. Four of these are shown in

Fig. 11-29. The types shown in Fig. 11-29c and d are envelope and radial presentations, described in Secs. 11-6 and 11-16, respectively. The presentation of Fig. 11-29a is produced by applying signals of equal phase but various amplitudes to the two sets of deflecting plates. This method is particularly suited to the instantaneous cathode-ray direction finders, which determine direction by comparing the amplitudes of the signal measured in two

crossed-loop antennas. The presentation of Fig. 11-29b is merely a linear presentation of the pattern of Fig. 11-29c. A receiver connected to a rotating directional antenna has its output connected

to the vertical set of deflecting plates of the cathode-ray tube. The horizontal axis is obtained by a sweep circuit that drives the trace across the screen for every revolution of the antenna.

The type of presentation shown in Fig. 11-29a is not widely used because the direction finders to which it is most suited, the so-called *cathode-ray direction finders*, use dual-channel receivers. Since the accuracy of the system depends upon the relative output of the two channels, the r-f, oscillator, and i-f sections of the two receivers must be made to track throughout the frequency range. The present state of dual-channel receiver development tends to exclude this method of presentation for accurate radar direction-finding applications.

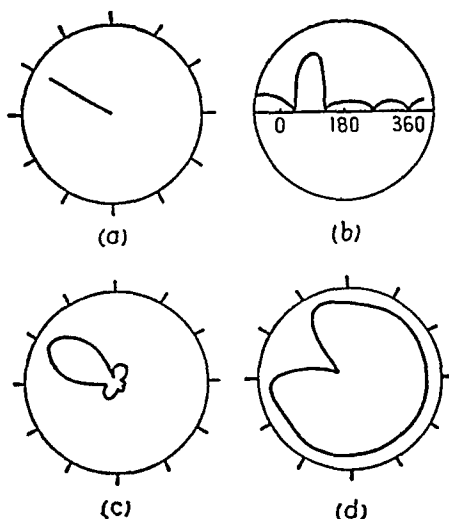


FIG. 11-29.—Various methods for cathode-ray-tube indication: (a) line indicator; (b) linear field-strength pattern; (c) envelope-type field-strength pattern; (d) radial pattern.

The presentations shown in Figs. 11-29*b*, *c*, and *d* are all suited to direction finders with rotating antenna systems. If the antenna has one direction of maximum response and comparatively small minor-lobe response, the question arises as to which of the three methods of presentation is most easily and accurately read. Most observers can bisect the lobe of Fig. 11-29*c* with greater accuracy than that of Fig. 11-29*b*. The radial method of Fig. 11-29*d* sharpens the pattern toward the center of the screen, but tends to spread it out around the edge of the tube. Although the central part can be easily divided, the present production tolerances on radial-type cathode-ray tubes make the deflection errors near the center of the screen unreasonably large.

If a presentation similar to that of Fig. 11-29*c* is used with a null-type direction finder, *i.e.*, if the point of maximum deflection corresponds to the minimum signal, instead of the maximum signal, the direction of the maximum deflection is not so reliable, because the noise-interference level may be above the minimum signal level. In this case the maximum error appears at the tip of the lobe, and the maximum bearing accuracy is obtained not by taking the direction of maximum deflection, but rather by bisecting the lobe, using the portions of the pattern closest to the center of the screen as a guide.

11-20. Reference Bearing.—For land-based direction finders the bearing information is generally presented in terms of degrees measured from the true north. On aircraft the presentation is usually relative, and the bearing is given with respect to the heading or longitudinal axis of the plane. For shipborne installations the presentation may be with respect to either the ship's heading or the true north. If true bearings are desired for shipborne or airborne direction finders the simplest method of allowing for calibration errors is to present the information in terms of relative bearings and to provide supplementary means of converting these to the true bearings. The direction-finder deviations depend on the antenna location and may or may not be small enough to allow a direct presentation of true bearings.

11-21. Instrument Errors.—It is interesting to consider the instrument errors of a typical indicator unit, such as the indicator unit described in Sec. 11-7 and shown schematically in Fig. 11-9. Although these errors do not hold for all the indicators described, the order of accuracy is approximately the same. The total instrument error is the sum of the errors of the various components that comprise the system and can be attributed chiefly to the following components:

1. *Synchros.*—The normal production tolerance on the electrical accuracy of a synchro is of the order of 0.5 to 1 deg, depending on the synchros used. When two synchros are used in a generator-follower combination, a variation of approxi-

mately plus or minus 1 deg should be expected. This error is independent of other portions of the servo system as it must be assumed that the follower synchro has negligible friction.

2. *Scanning Capacitor*.—Two types of errors arise in the scanning capacitor. The first is electrical, the result of the misalignment of plates due to production tolerances and mechanical handling. Normal errors for such a capacitor are generally of the order of plus or minus 2 deg. Because the four stator plates are independent, the error generally takes the form of a misalignment in one or more quadrants.

The second error in the scanning capacitor is mechanical and can be attributed to the rolling friction of the bearings and the windage encountered as the speed of rotation increases. This latter error differs from the electrical error in that it can change, not only with the aging of the bearings, but also with the speed of antenna rotation. Its effect becomes evident when it is observed that the apparent direction of the signal source shifts as the speed of the antenna is changed. This error may amount to 3 or 4 deg for various capacitors.

3. *Cathode-ray Tube*.—A third instrument error arises in the alignment of the deflecting plates of the cathode-ray tube. Production tolerances for normal cathode-ray tubes allow 90 deg plus or minus 3 deg between the two sets of deflecting plates. The instrument error caused by such a misalignment is symmetrical in that the errors are common to opposite quadrants. Thus bearings that lag between 0 and 90 deg will lag by a similar amount for bearings between 180 and 270 deg.

4. *Trimmer Capacitors*.—A number of difficulties confront the maintenance man in the alignment of trimmers. For the system of Fig. 11-9 the procedure requires him to adjust the capacitors while observing the luminous trace on the cathode-ray tube. In general, the attainable accuracy is not more than plus or minus 1 deg unless special equipment and techniques not generally adaptable to field maintenance are used.

Summary.—The errors outlined above may not apply specifically to all the indicator units described in this chapter, but the nature of the problems and the order of magnitude of the errors are typical. In using radar or radio direction finders for high-frequency applications, it can be said that the instrument error is generally commensurate with that of the antenna array. It usually limits the accuracy of the system to the order of plus or minus 5 deg.

CHAPTER 12

HOMING SYSTEMS

By J. W. CHRISTENSEN

12-1. Introduction to Homing.—Homing is a specialized form of radio direction finding. With direction-finding equipment, bearings may be taken with any arbitrary heading of the vehicle; while in homing, the bearing is taken by orienting the vehicle until its heading is toward the source of radiation. Greater bearing accuracy can, in general, be

obtained with homing than with direction finding because of the narrower sector over which accuracy is required.

A typical homing system, as installed on an aircraft, is shown in Fig. 12-1. Although many variations are possible, the essential components and principles remain the same. The signals received by two or more directional antennas on the aircraft are compared, and this information is presented to the pilot in such a manner that he can correct his

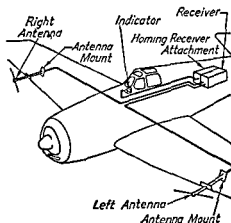


FIG. 12-1.—Azimuth homing system for horizontal polarization.

Homing systems have been built to provide directional guidance in the horizontal plane (azimuth homing), in the vertical plane (elevation homing), or in both planes. Elevation homing has proved to be less useful and technically more difficult.

In practice, high-frequency homing systems consist of the antenna system, a commutating system, and an indicator. These are used in conjunction with a standard u-h-f receiver. A block diagram of a typical azimuth homing attachment is shown in Fig. 12-2. As the figure shows, the homing attachment provides, by means of a switching arrangement, a method whereby a single receiver can in effect yield an indication of the relative voltage induced in two separate antennas by a single transmitter. Approximately half the time, for example, the receiver is connected so as to activate the entire circuit or channel from the left antenna to the indicator. The desired end product from this system is an indication on the

cross-pointer meter which will keep the pilot informed of the correctness of his course. In the equipment shown, the vertical needle is caused to swing right or left to indicate errors in heading, while the horizontal needle is used to indicate the integrated signal strength and thus inform the pilot of his approach to the signal source. For homing in both azimuth and elevation, two more channels would have to be added, and the horizontal pointer of the cross-pointer indicator could then indicate any errors in the flight attitude.

The following sections describe the results of work at the Radio Research Laboratory on homing systems operating in the frequency range

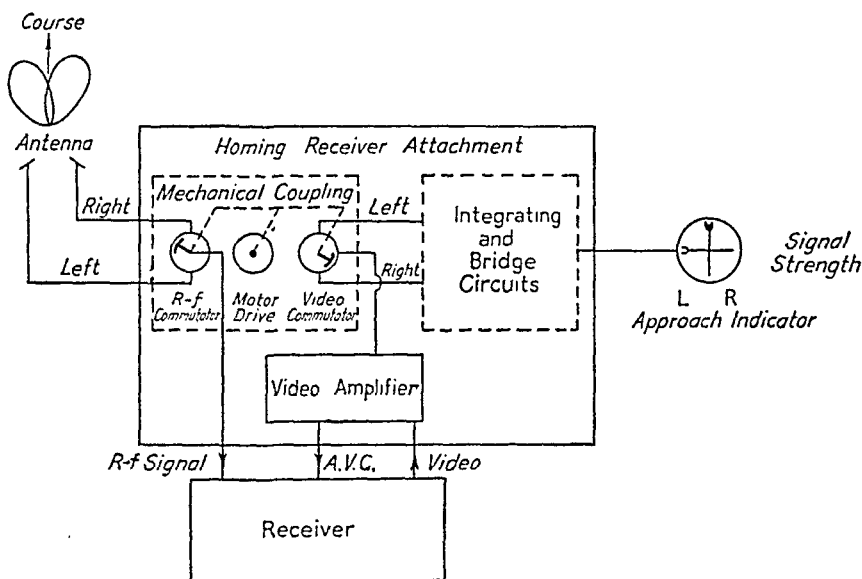


FIG. 12-2.—Azimuth homing system. Block diagram.

above 100 Mc. These are of interest not only because of the high operating frequencies, but because of the requirements met for operation over a wide band of frequencies. Various design features are also of interest. The homing systems described here are all of the lobe-switching type.

12-2. Antennas for Very-high-frequency Homing. Basic Considerations.—The antenna system is the heart of any homing device. The design of suitable antennas represented the principal new problem faced in extending the frequency range of homing devices to the u-h-f region.

Ambiguous Courses.—A balance or on-course indication is given to the pilot by the homing system whenever his heading is such that equal signals are induced in each of the antennas of a pair. As shown in Fig. 12-3a, there may be several of these crossover points at which the radiation patterns of the two antennas intersect. Crossover points must,

of course, occur in pairs, although one point of a pair may be in or near a direction of zero sensitivity. An ideal homing pattern would have only one crossover point, and that dead ahead of the aircraft.

In practice, most homing antenna systems have, in addition to the true crossover point, one or more ambiguous crossover points. Use of the ambiguous ones by the pilot should be prevented. The determination of the ambiguous crossover points is called *sensing*, and may be accomplished by one or more of the following means. (1) One-half of the crossover points give reversed reading and so can readily be identified on the pilot's indicator. For example, in an azimuth homing system, if the aircraft deviates to the right, correct sense indicates that the pilot should turn left, while reverse sense indicates that he should turn even more to the right. By the use of this principle, complete sense can be obtained with the double cardioid patterns often used at lower frequencies. (2) If the sensitivity forward can be increased relative to the sensitivity

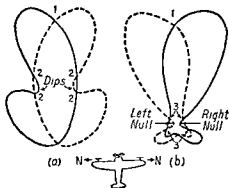


FIG 12-3 — Ambiguous crossover points due to dips and nulls (1) true crossover point, (2) ambiguous crossover point due to dips; (3) crossover points due to nulls; N → direction of null

backward, the forward and backward groups of crossover points can be separated by means of a signal-strength indicator. If all crossover points except the true one can then be moved to the backward quadrant, a signal-strength indicator will yield complete sense.

Deviation of the true crossover point from dead ahead obviously occurs when the left-right or up-down antenna pattern pairs are not symmetrical. Causes of such dissymmetry may be (1) unsymmetrically mounted antennas, (2) electrically unbalanced antenna

cables, (3) a reflecting structure near one antenna, (4) dissymmetry in electrical ground paths and metal surfaces surrounding antennas.

Much of the work discussed in this chapter was originated in an effort to modify the antenna patterns to remove ambiguous crossover points, such as shown in Fig. 12-3. In particular, with wing-mounted antennas such ambiguous crossover points often occur due either to *dips* (sharp decreases in sensitivity over limited forward angles) in the antenna pattern, or to *nulls* in the direction of the ends of the wings. Dips are sometimes difficult to overcome, especially when the antennas are mounted on the wings, but may often be eliminated by placing the antennas further outboard or by increasing the distance between the antenna and the leading edge of the wing. Nulls toward the ends of the

wings may often be eliminated or moved backward by mounting or bending the antennas so that the outboard end of each antenna is closer than the inboard end to the leading edge of the wing, and by mounting antennas further outboard. A homing antenna system developed on the basis of such considerations is discussed in some detail later.

Sensitivity vs. Definition.—A compromise is usually found necessary between sensitivity and directional accuracy. If the two antenna patterns of a pair are arranged as shown in Fig. 12-4a, the true crossover point appears in the direction relative to the antenna patterns which yields high sensitivity. However, the change of relative signal received from the two antennas is not great with changing aircraft heading. Therefore the directional definition is low. When the antenna patterns are arranged as in Fig. 12-4b, angular definition is high but sensitivity is low.

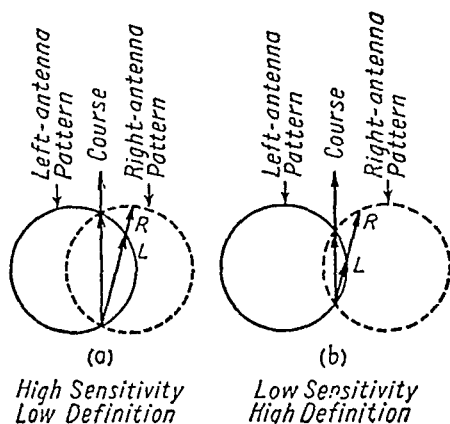


FIG. 12-4.—Sensitivity vs. definition of antenna patterns: (a) high sensitivity; low definition; (b) low sensitivity; high definition.

Polarization.—The effects of the polarization of the incident wave on the performance of a homing antenna system have been found to be marked. The homing systems described in the following pages will operate satisfactorily only on one polarization. The difficulties of maintaining proper balance under all conditions of polarization have not yet been surmounted at these frequencies. In the 100-to-600-Mc band, a separate pair of antennas is, in general, required for each angular definition on each polarization. Simple linear elements or stubs are usually sufficient to give azimuth indication on horizontal polarization or elevation information on vertical polarization. Phased arrays, a linear element with reflector, or stub-reflector combinations are usually necessary to yield azimuth information on vertical polarization or elevation information on horizontal polarization. Various examples of these types of homing antennas are discussed later.

12-3. Pattern Measurement.—To facilitate rapid determination of homing-antenna patterns, two pattern-measuring arrangements were constructed with which both azimuth and elevation patterns could be measured on the ground. One of these is illustrated in Fig. 12-5. A test transmitter is mounted in a truck that can be driven around the aircraft while a calibrated receiver and output indicator are connected to the homing antenna of the aircraft. The radiation pattern of the antenna

under test may be then readily determined. It is essential that no buildings or other reflecting objects be near the testing area.

To assist in the semi cut-and-try process of design, it was found convenient to mount the antennas under test on flexible gooseneck tubing so that their spacing and orientation on the aircraft could readily be changed. For such tests on the ground, it was not found necessary to make permanent installation of the antenna, provided that the contours of the aircraft were not markedly changed by the test equipment.

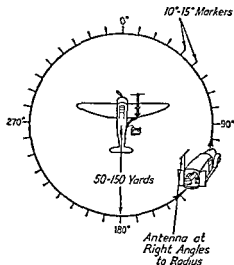


FIG. 12-5.—Over-all azimuth pattern-measuring arrangement.

The transmitting antenna on the truck may be a single dipole mounted so as to give the desired polarization. It is extremely important that the dipole be balanced to ground so that the polarization of the signal radiated cannot have both vertically and horizontally polarized components. Signals of mixed polarization will produce misleading results in such a test.

Patterns for elevation homing with wing-mounted antennas can be obtained through the use of a full-sized metal aircraft wing mounted with its long axis vertical

in such a manner that it may easily be rotated. In this case, the transmitter remains stationary, and the receiving antenna is rotated. The same equipment may then be used as for the determination of over-all azimuth pattern.

12-4. Classification of Antenna Systems.—A distinction may be made between two general types of homing antennas: a self-complete type, the operation of which is influenced by, but not dependent upon, the aircraft structure; and a structure-dependent type, the operation of which is directly dependent upon the aircraft structure on which it is mounted. Tests on self-complete antennas indicate that unidirectional patterns having high forward-to-backward sensitivity ratios can be obtained with Yagi or similar-type arrays, but only over approximately a 10 per cent frequency range. Where greater frequency ranges such as 1.5-to-1 or 2-to-1 frequency ratios are desired, simple dipole and wide-band-Adcock combinations are satisfactory. Self-complete antennas of this type have the advantages of operation through a fairly broad frequency range and on practically any type of structure, when mounted sufficiently clear of it. They have the disadvantages of having ambiguous crossover points and of requiring considerable mounting space.

Structure-dependent antennas, such as simple stubs or dipoles, are mounted at locations where metallic surfaces of the aircraft can be used as shields or reflectors to produce proper patterns. In general, such antenna systems yield satisfactory patterns over frequency ranges of 2 to 1 or 2.5 to 1. Detailed examples of both types of antennas are given in a later section.

For proper operation of homing antennas of the dipole type, it is always desirable and usually necessary to operate them balanced to ground. Such balance is best achieved by the use of a balun. Correct balancing becomes increasingly important as frequency ranges are increased. A detailed description of the operation of such balanced transformers is given in Chap. 3.

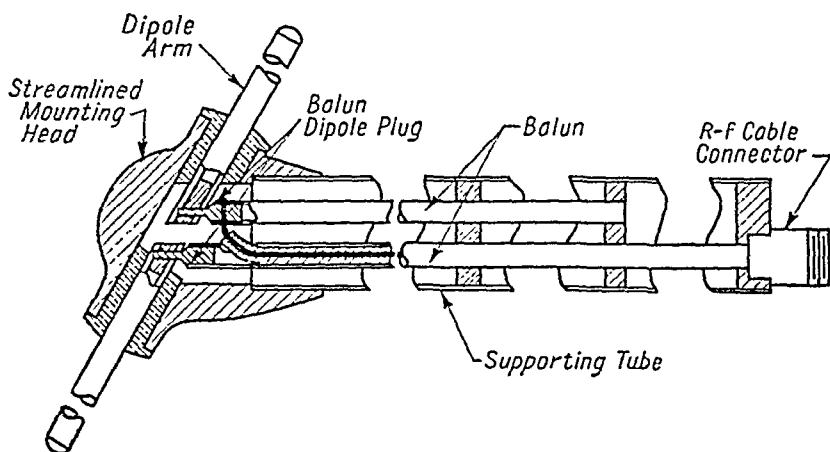


FIG. 12-6.—Wing dipole antenna.

It is important that the cables connecting the antennas with a lobe-switching equipment be of identical electrical length, since standing waves present on these cables may cause directional errors if the cables, and hence the waves, are not identical.

12-5. Antenna Systems for Azimuth Homing. *Double-dipole Wing Antenna for Horizontal Polarization.*—The homing antenna found most suitable for use on horizontally polarized signals is a structure-dependent system consisting of a pair of dipoles, one mounted on the leading edge of each wing of a metal aircraft, as shown in Fig. 12-1. A cross-sectional view of one of the antennas and its associated balun is shown in Fig. 12-6. Polar sensitivity curves for the over-all system are shown in Figs. 12-7*a* and *b* at frequencies of 150 and 300Mc, respectively. At intermediate frequencies the patterns are equally good. The dipole is $\frac{1}{2}$ wavelength long at 180 Mc.

In this antenna system, the wings and nose of the metal aircraft are used as reflectors to assist in obtaining the desired polar patterns. Shielding by the wings is largely responsible for the high forward-to-backward

sensitivity ratio of this system, which is of the order of 20 to 25 db, depending on frequency. Shielding by the nose, combined with the sweepback of the wings, causes most of the right-left displacement of the starboard and port antenna patterns. The 30-deg angle between the

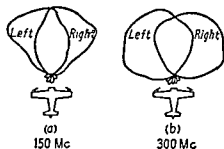


FIG. 12-7.—Wing dipole antenna patterns.

antennas and the leading edge of the wing tends to fill in dips in the patterns toward the wing tips that would otherwise be present and causes a slight increase in the right-left lobe displacement.

It has been found that, if the antenna position is too near the fuselage, sensitivity forward becomes less, and secondary dips appear in the patterns at approximately 60 deg

and 300 deg. These dips cause serious ambiguous crossover points. When the mounting position is too near the wing tips, on the other hand, the right-left displacement becomes insufficient and the forward-to-backward sensitivity ratio decreases. This antenna is the most satisfactory azimuth homing system developed for horizontal polarization since it has only one crossover point forward, and all back crossover points are 20 to 25 db less sensitive than the forward one. With this antenna, pilots find it relatively simple to separate the forward and backward crossover points either by using a signal-strength meter, or simply by noticing the jumbled nature of the backward crossover points.

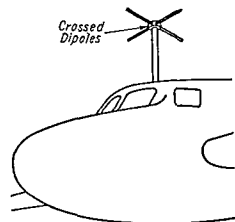


FIG. 12-8.—Crossed dipoles mounted.

Crossed-dipole Antenna.—An antenna of the self-complete type that yields azimuth information on horizontally polarized signals is shown mounted above an aircraft in Fig. 12-8. The dipoles are mounted at right angles to each other in a horizontal plane, each dipole making an angle of 45 deg with the axis of the aircraft. Patterns with this antenna so mounted are shown in Fig. 12-9. It can be seen that the effect of the aircraft is to distort the patterns from their normal shape, yet the crossover points were found to remain within 4 deg of the correct position over a frequency range of about 1.6 to 1. An auxiliary sensing antenna having a good forward-backward sensitivity ratio is required in order to

enable the pilot to distinguish between the 0-deg correct bearing and the 180-deg ambiguous bearing.

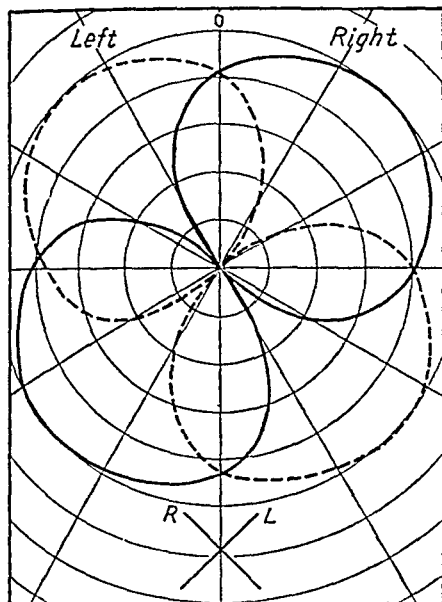


FIG. 12-9.—Patterns of crossed-dipole antenna mounted above plane surface, 175 Mc.

Triple-stub Antenna for Vertical Polarization.—The antennas shown in Fig. 12-10 proved very satisfactory for providing azimuth homing on vertically polarized signals. The antenna consists of two identical stubs, installed side by side and normal to the bottom surface of the fuselage of a metal aircraft. Between the stubs a heavy metal

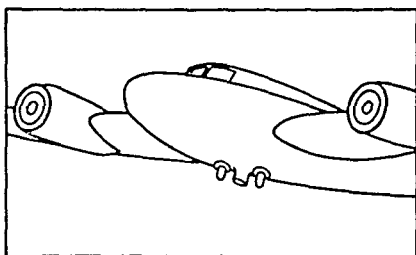


FIG. 12-10.—Triple-stub antenna.

streamlined strut is grounded to the fuselage. This strut is at least 10 per cent longer than $\frac{1}{4}$ wavelength at the lowest frequency for which the system is to be used. The stubs are about $\frac{1}{4}$ wavelength long and are spaced about $\frac{1}{4}$ wavelength from the streamlined strut at the center frequency. The resulting patterns at one frequency are shown in Fig. 12-11.

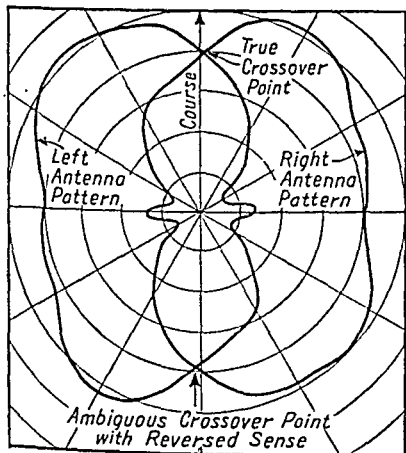


FIG. 12-11.—Triple-stub antenna patterns at 580 Mc.

This antenna is highly satisfactory for yielding azimuth information on vertically polarized signals, since it has good sensitivity, good definition, and relatively wide frequency range. Crossover points remain within a few degrees of the correct directions over a 1.5-to-1 frequency range. The greatest application is likely to be for

use on aircraft at frequencies above about 200 Mc since, at lower frequencies, the center strut must be quite large.

12-6. Antennas for Elevation Homing. *Double-dipole Wing Antenna for Horizontal Polarization.*—A double-dipole wing-antenna system that yields elevation information on horizontally polarized signals is shown in Fig. 12-12. This antenna consists of two horizontal dipoles mounted one above the other and slightly forward of the leading edge of the wing. The wing acts as a reflector, causing vertical displacement of the patterns

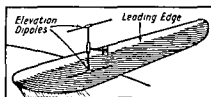


FIG 12-12.—Double-dipole-wing elevation antenna.

and giving a considerable forward-backward sensitivity ratio. A typical pattern in the vertical plane for this system is shown in Fig. 12-13. Because of the unsymmetrical cross section of the wing, elevation patterns vary with frequency. It is usually possible, however, to find a position for the antennas where the patterns are satisfactory for homing and where the vertical crossover point remains within plus or minus 3 deg of the aircraft axis over a frequency range of approximately 1.5 to 1.

Crossed-Adcock Antenna for Horizontal Polarization.—An elevation-homing antenna system of the self-complete type for use on horizontal polarization may also be constructed from two H-type Adcock antennas with horizontal elements, one mounted on each side of the nose of a multimotored aircraft. One antenna can be set so that its pattern maximum is 45 deg above the horizontal plane, and the other with its pattern maximum at 45 deg below the horizontal plane. Such an antenna has excellent definition and is usable over approximately a 2-to-1 frequency range, but has the disadvantage of ambiguous crossover points. The backward ambiguous crossover point is especially serious. Other disadvantages of this type of antenna are the large mounting space required and severe aerodynamical problems.

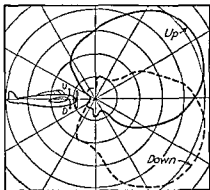


FIG. 12-13.—Double-dipole-wing elevation-antenna patterns at 175 Mc.

Universal Homing Antenna.—A combination of the crossed-dipole and the crossed-Adcock antenna is shown in Fig. 12-14. The dipoles and Adcocks are all mounted on a central hub, carried by a supporting tube mounted parallel to the axis of the aircraft. Rotating the supporting tube through 90 deg will orient all the antenna elements for use on either horizontal or vertical polarization. For horizontal polarization

the crossed dipoles give azimuthal information and the crossed Adcocks yield elevation information, conversely for vertical polarization. A large circular shielding plate is mounted on the supporting tube a short distance from the hub. This increases the forward-backward sensitivity ratio and decreases the effect of near-by objects behind the antenna mount. The best patterns from this type of antenna are obtained when the antennas of the Adcocks are cut to $\frac{1}{2}$ wavelength and spaced $\frac{1}{2}$ wavelength apart and the dipoles are cut about 20 per cent longer than $\frac{1}{2}$ wavelength. The presence of ambiguous crossover points both in the vertical and horizontal planes limits the use of this type of antenna to accurate homing on a course that has previously been approximately determined by other methods.

12-7. Commutators.—The r-f input to the receiver and the video-frequency output both must be commutated rapidly and in synchronism. The proper operation of a commutating system is, therefore, essential to the operation of a lobe-switching homing device. Both mechanical and electronic commutators have been built and have given satisfactory performance.

Mechanical r-f commutators for this frequency range have been found to perform better if the switching contact is made through a capacitance. Two types of switching transients are produced by metal-to-metal contacts. One type of transient is due to the switching of static charges on the antenna system. The effect of this transient may be easily eliminated by placing a small choke between the inner and outer conductors of the antenna coaxial line to prevent the build-up of static charges. A second and much more serious type of transient is due to a frequency or amplitude shift of the receiver local oscillator, whenever the antenna input circuit is switched. This is caused by a coupling of energy from the local oscillator into the antenna input line through the mixer and r-f tuning sections of the receiver. Switching the antenna circuit changes the load of the local oscillator and thereby causes the above shift. This type of transient varies with the receiver used, the length of coaxial line involved, and the frequency. The peak amplitude of these switching pulses varies from 100 to 500 μ v. This means that a metal-to-metal contact switching system seriously limits the useful sensitivity of the receiver.

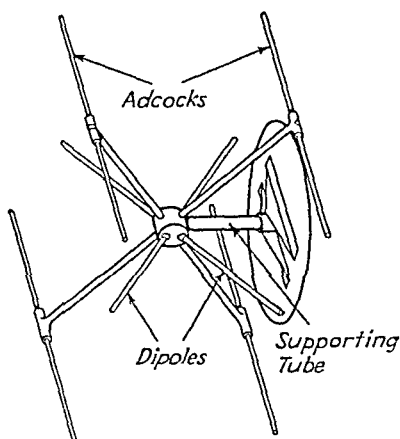


FIG. 12-14.—Universal homing antenna.

12-8. Mechanical Radio-frequency Commutators.—Rotary commutators of the automobile high-voltage ignition-distributor type have been found to be the most satisfactory of the mechanical r-f commutators. In such a commutator, input segments connected to the inner conductors of the antenna-input cables and an output disk connected to the inner conductor of the cable to the receiver are embedded in a low-loss phenolic stator plate. A shielding ring is provided between the center disk and the input segments to reduce capacitance coupling and cross talk and to improve the impedance match. Rotor shoes of stainless steel make light sliding contact with the phenolic stator plate. The input segments and the output disk are undercut 0.0015 in. below the surface of the stator plate. Hence, for each channel, two air-dielectric capacitors are in series when the commutator makes contact. The rotor, which is rotated about 900 rpm, then couples in turn, through these capacitances, the various input segments to the output disk. Such a commutator may be made to cause less than 6-db insertion loss in a 50-ohm coaxial cable over a frequency range of about 5 to 1 and less than 12-db loss over a frequency range of about 10 to 1. Such commutators can be constructed that operate with less than 6-db loss up to frequencies as high as 1300 Mc. One life test on such an air-capacitor-type commutator showed it to be in good working condition after 3000 hr of continuous operation.

Somewhat lower losses can be obtained by the use of titanium dioxide instead of air as the dielectric in a mechanical commutator similar to the one discussed above. In such a commutator, the input segments and the output contact protrude above the stator plate, and a titanium dioxide coating on the rotor makes direct rubbing contact with these segments. In a unit of this type, it is possible to use polystyrene as a stator material. Polystyrene stator installation cannot be used in the air-dielectric type commutator because it would be rapidly worn away by the rotor segments. A titanium dioxide commutator has been made to operate with less than 7-db loss to frequencies as low as 20 Mc.

The fact that titanium dioxide is very hard and tends to cut away the material on which it rubs limits the life of such a commutator. Under good conditions, a life of about 200 hr may be expected.

Single-pole double-throw coaxial r-f relays have also been used for r-f commutation in homing systems. These relays have the advantage of low loss and low standing-wave ratio. However, they suffer from the transients generated and from bouncing of the contacts at the higher switching rates. Capacitive switching may be obtained in such a relay by anodizing the relay contact points. This eliminates the switching transients, but such a relay has a relatively short life. For low switching rates as used in an aural indication of the A-N type, such relays are satisfactory.

12-9. Mechanical Video Commutator.—The mechanical construction of a video commutator may be similar to that used in the r-f commutator just described, except that direct contact is used. Figure 12-15 shows the essential components of a two-channel video commutator.

Because of the low current switched by the video commutator of the homing system, two serious problems of commutator film are likely to be encountered: one with film built up between the segments and the brushes, thereby increasing the commutator series resistance; and the other with the film built up between adjacent segments, thereby causing cross talk between adjacent channels. Commutators employing brass segments and ordinary carbon or graphite brushes are unsatisfactory and cannot be relied upon for more than 5 to 10 hr of operation.

A combination of sterling silver segments, brushes 90 per cent silver and 10 per cent graphite, and brush pressures of between 5 and 8 lb per sq in. is about optimum for keeping series resistance low over long operating periods. Small slots, approximately 0.02 in. wide, at the ends of each segment may be used to break up the intersegment film and thereby prevent cross talk between adjacent channels. Using this combination, life tests on several commutators indicate a useful life in excess of 3000 hr, during which the series resistance did not become greater than 0.80 ohm.

The r-f and video commutators described above are usually mounted in a pair on a small casting and driven in synchronism by a small d-c motor.

12-10. Electronic Radio-frequency and Video Commutator.—To eliminate all moving parts, an electronic commutator has been developed which is composed of r-f and video gates controlled by an electronic switch. A block diagram of this system as applied to simultaneous azimuth and elevation homing is shown in Fig. 12-16. The r-f gate consists of four grounded-grid 6J4 amplifiers capable of operating over a very wide frequency range. The plate voltage of each of these gate tubes is switched on and off, as required, by the electronic switch. The video gate consists of four triode grid-bias-controlled rectifiers of conventional type, also controlled by the electronic switch. The electronic switch is a four-channel Eccles-Jordan circuit triggered by a blocking oscillator. Figure 12-17 shows the details of this circuit.

Assume that triode V_1 is conducting space current; this causes a volt-

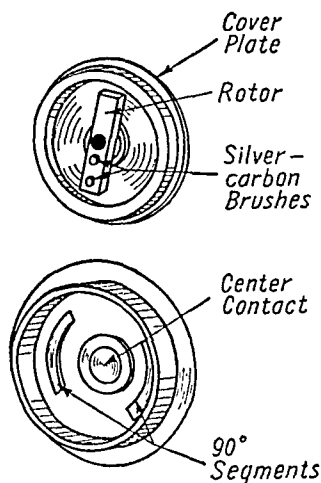


FIG. 12-15.—Two-channel video-mechanical commutator.

age drop in the corresponding plate-load resistor. Since the grids of the other three triode elements are connected through identical voltage-dividing resistance networks to the plate of V_1 , these grids are made less positive with respect to ground (more negative with respect to their

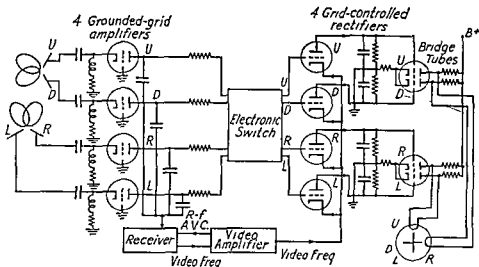


FIG. 12-16.—Block diagram of electronic homing attachment.

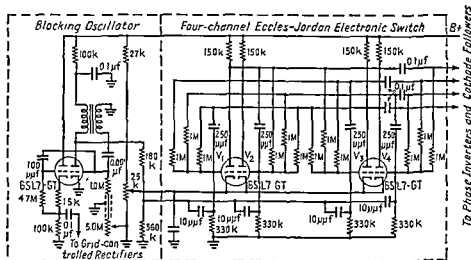


FIG. 12-17.—Schematic diagram of two- or four-channel electronic homing attachment.

cathodes) than otherwise. This holds V_2 , V_3 , and V_4 to cutoff. Now, if a large negative pulse from the blocking oscillator is applied to all the grids in parallel, V_1 will suddenly be driven to cutoff; the other tubes will be unaffected. This causes a rapid rise in the plate voltage of V_1 , which is applied only to the grid of V_2 through the coupling capacitor.

The amplified positive pulse applied to V_2 overrides the negative pulse from the blocking oscillator, and V_2 conducts. This stable situation corresponds to that explained above for V_1 conducting. The next negative pulse will then transfer control to V_3 , and so on.

The output of each channel of the Eccles-Jordan circuit is passed through pulse-inverting and shaping tubes to cathode followers. The low-impedance output of the cathode-followers controls the r-f and video gates in synchronism.

The loss through a typical switch of this type is approximately 6 db at 10 Mc, 7 db at 200 Mc, 8 db at 400 Mc, 10 db at 500 Mc, and 16 db at 900 Mc.

Comparative tests between the electronic and mechanical types of commutator indicate that both are satisfactory for use with homing systems. The air-dielectric mechanical commutators are superior for frequencies above approximately 300 or 400 Mc, while the electronic type is better for frequencies below this value. The titanium dioxide mechanical commutator yields low-loss results as low as 20 Mc but has mechanical difficulties. The mechanical commutators are limited to switching rates of the order of 15 to 20 per sec, while the electronic commutator may, of course, switch at any desired rate.

12-11. Indicators.—The relative merits of the several methods of presenting homing information to the pilot were investigated. Three systems seem most promising.

1. An oscilloscope indicator, in which the four right-left up-down deflection plates are connected to the rectified outputs of the corresponding homing channels, may be used for azimuth and elevation homing, since polar deflection of the oscilloscope spot indicates the direction of the transmitter being homed on. This system has only fair definition due to the relatively high deflection voltages required, is heavy, has a large number of controls, and has a tendency to develop "burned" spots on the screen. It is likely, however, that an improved system of this type could be used to advantage in discriminating between sources having the same r-f frequencies but differing in modulation characteristics.

2. A cross-pointer meter having two movements in a single case with the vertical needle indicating azimuth direction and the horizontal needle indicating elevation information (or relative signal strength for an azimuth-only homing system) is an excellent indicator due to its light weight, simplicity, ease of reading, and low power requirements.

3. An A-N aural keyed system also works very satisfactorily as the indicator for two-channel operation. In this system, the right and left antennas are keyed in an interlocking A-N in the usual radio-range fashion by the r-f commutator.

From laboratory and flight tests comparing the relative merits of the A-N aural system and the cross-pointer indicator when used with similar homing equipment the following results were obtained: (1) A-N aural

presentation was found to be unsatisfactory for use with a combination azimuth and elevation system because of difficulty in interpreting a signal keyed four ways; (2) preliminary psychological tests indicate the definition of the meter system to be approximately 1.5 times the definition of the A-N system; (3) flight tests indicate that meter presentation yields two to three times the definition of A-N presentation, probably because of the noise in the aircraft and the inability of the pilots to concentrate on the A-N signals; (4) most pilots prefer meter to A-N because of the fatigue attendant upon concentrating on the A-N signals for very long periods of time; (5) A-N presentation allows some resolution of stations having the same carrier frequency and differing only in modulation; meter presentation allows no resolution.

12-12. Miscellaneous Considerations. A-v-c Circuits.—In homing on a radio transmitter, the received signal strength varies between very wide limits. If the receiving system is to operate with adequate sensitivity and without overloading, a very effective automatic volume control is necessary. At the same time, the time constant of the a-v-c circuit must be sufficiently long to keep the gain of the receiving system essentially constant over several rotations of the r-f and video commutators. With the standard receivers used, the automatic volume control normally present is insufficient and the time constant is too short. It is, therefore, necessary to provide from the homing attachment to the receiver an a-v-c voltage of sufficient amplitude and of long time constant.

Two stages of video amplification before a-v-c rectification and an a-v-c time constant of approximately 1 sec were found satisfactory for the particular receivers used. With this combination and a typical receiver, satisfactory operation can be obtained with r-f inputs varying from 20 μ v to 1.0 volt.

Integrating and Bridge Circuits.—The integrating and bridge circuits used in v-h-f homing equipment are not essentially different from those used in earlier low-frequency systems. These circuits integrate the signals coming from a pair of channels, compare them as to intensity, and develop an indicator current, the magnitude of which is roughly proportional to the difference of intensity between the two incoming signals. As Fig. 12-16 shows, the bridge circuit operates from the rectified negative-peak output of the video amplifier in the receiver. The time constant of the RC integrating network shown in the grid circuit of the bridge tubes is chosen to be long (0.1 to 0.5 sec) compared with the switching period. This ensures a steady indication on the cross-pointer meter. If the airplane is off course, different r-f voltages will be induced in the two antennas of a pair, and hence different d-c grid voltages will be developed on the two halves of, say, the right-left bridge tube. The resulting different plate currents will cause a potential difference to

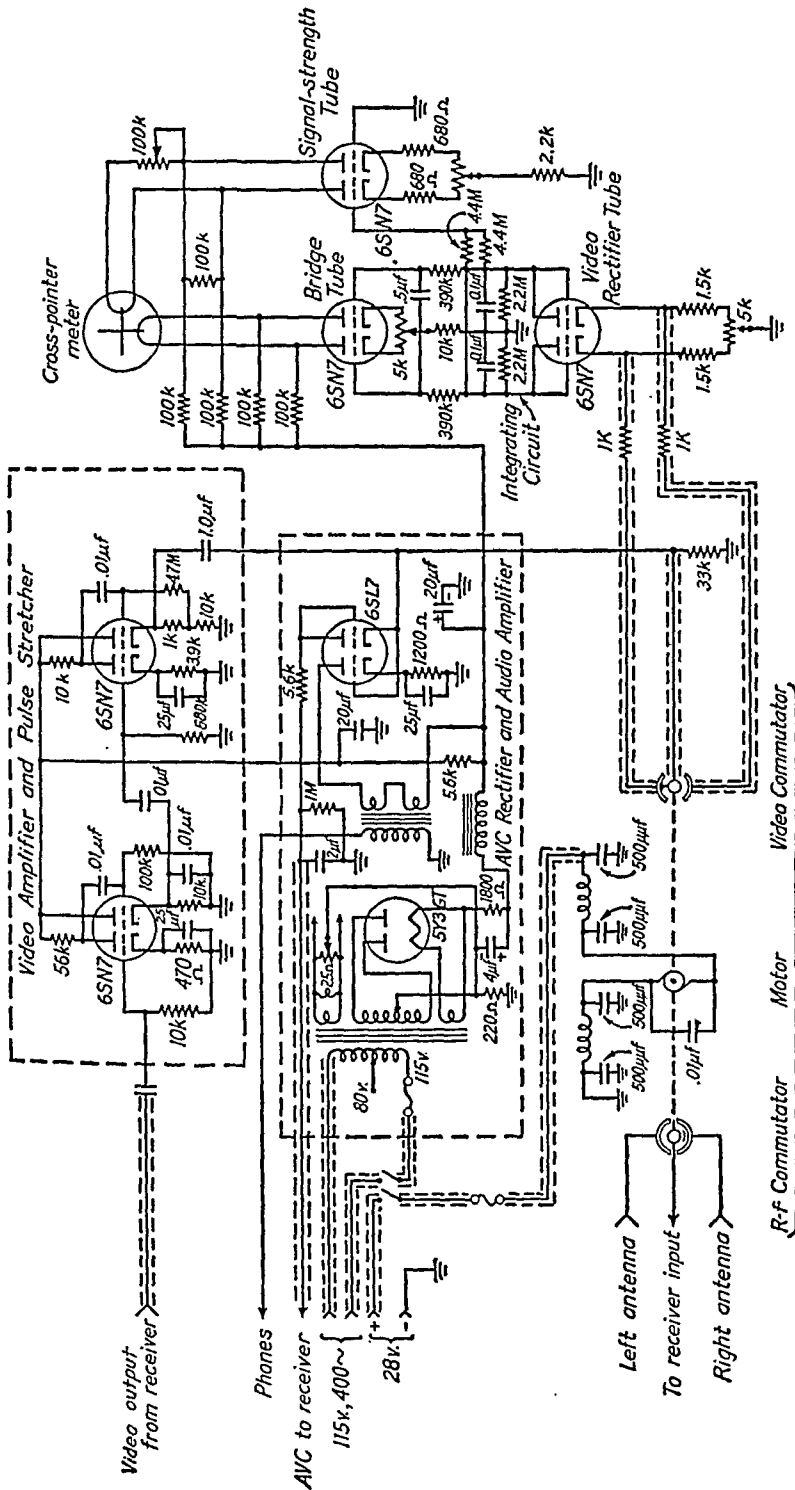


Fig. 12-18.—Schematic diagram of two-channel mechanical homing attachment.

appear between the two plates, the sign of which will depend on the direction of the aircraft heading error. The pilot's indicator is connected to read this plate-to-plate voltage. As shown in Fig. 12-18 (the schematic diagram of a widely used system), the circuit is composed of a double rectifier, double integrating circuits, a commutator-ripple filter, and a double bridge tube feeding the indicator. In this particular system, an additional tube is used to actuate the horizontal-needle element of the cross-pointer indicator in accordance with the average signal strength (considering *both* right and left channels). Information on signal strength aids the pilot in distinguishing ambiguous crossover points and in determining the approximate distance to the signal source.

12-13. Performance.—The double-dipole-wing antenna system for azimuth homing on horizontally polarized signals proved to be the most useful. Extensive tests with this type of equipment indicate that it is possible consistently to obtain bearing accuracies of the order of ± 5 -deg error over a 2-to-1 frequency range. The errors are not greater than ± 3 deg at most frequencies. Accuracy can, in general, be made higher than the pilot can effectively use in controlling his aircraft.

Tests of combined azimuth-elevation homing systems on a ground-based source indicate azimuth and elevation accuracy of the order of ± 5 deg. Tests against an airborne source showed poor accuracy in elevation due to the interference of direct and ground-reflected waves. As is to be expected, this interference produces oscillations of the elevation indicator above and below the true reading. While this oscillation makes homing in elevation more difficult on airborne sources, it is usually still possible to locate the source aircraft by noting the decreased amplitude of the oscillations as the source is approached.

CHAPTER 13

INTRODUCTION TO POWER GENERATION

BY J. F. BYRNE

13-1. Military Requirements as to Frequency and Power.—The following chapters dealing with the general subject of power generation are based upon the results of an intensive development program carried on during the war. This program was established to meet an urgent military requirement for modulated c-w transmitting equipment specifically designed to create a maximum of interference with the operation of unfriendly radar systems. The magnitude of the technical program can best be appreciated by noting that the useful radar spectrum extends from roughly 50 Mc to upward of 10,000 Mc and that an unfriendly radar system could well utilize any channel in this tremendous frequency range. Consequently, in order to meet the military requirements, transmitters were developed that were capable of being quickly tuned over wide frequency ranges, preferably with single-dial control, so that the useful radar spectrum could be guarded with the minimum number of different equipments and equipment types. A notable achievement in this direction was the development of equipment, later described, capable of being tuned and operated at any frequency in the range 300 to 2500 Mc, with efficiencies that, throughout this very wide range, generally exceed those commonly obtained prior to the war at the low-frequency end of the band.

A thorough study of the characteristics of radar systems, such as peak power, antenna gain, and scope-presentation methods, was undertaken in order to establish the transmitter power levels required to meet the various military applications. The general results of this work indicated that equipment ranging from a few watts to many kilowatts of output power could be effectively utilized. As pointed out above, a series of equipments are naturally required to cover this enormous frequency range, and from the standpoint of high-power v-h-f operation this requirement posed a very formidable tube problem; for it has been a natural condition that, as frequency is raised, physical dimensions of the tube generators decrease and heat dissipation is limited, resulting in a reduction of both power input and power output.

13-2. Tube Program.—The general considerations involving power levels required to accomplish the technical objectives resulted in the establishment of a very large vacuum-tube development program; for

prior to the war, power levels were limited to values of the order of 1 kw at 100 Mc and 25 watts at 500 Mc. Figure 13-1 shows the state of development prior and subsequent to the establishment of the program. An item, not shown on the chart, is a 10,000-watt magnetron operable at frequencies between 400 and 600 Mc. In general, magnetron development was somewhat more rapid than that of high-power triodes and tetrodes. From the point of view of the minimum development and production job, wide tuning ranges were obviously desirable. At the lower frequency end of the spectrum tubes were developed that employed external tuned circuits, and this class of tube, for more or less obvious

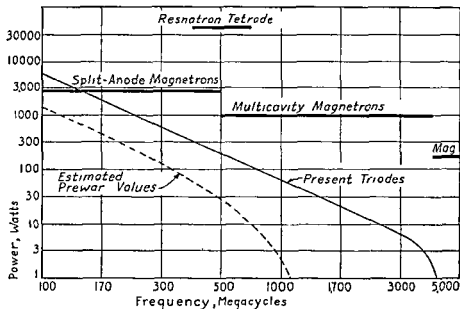


FIG. 13-1.—Maximum power levels resulting from the development program.

reasons, has the greatest frequency coverage in terms of the ratio of the high-frequency to the low-frequency limit, but not necessarily in terms of the total number of channels or megacycles covered. Table 13-1 lists a few tubes, their power output, tuning ratio, and the number of megacycles covered.

As is implied in Table 13-1, triodes, tetrodes, and magnetrons all found logical application in the program. The resnatron tetrode is probably one of the most outstanding electronic developments of the war period. At the time of its first successful test, power levels at 500 Mc were pegged at around 30 to 50 watts, and here in one package is a 30,000- to 50,000-watt generator—a thousandfold increase in one step. There is little doubt that the successful operation of the resnatron greatly accelerated all high-power high-frequency development and that it

points the way toward vastly increased domestic and industrial applications of high power in the frequency range up to 1000 Mc.

TABLE 13-1.—TYPICAL TUBE CHARACTERISTICS

Tube	Type	Tuned circuit	Frequency range, Mc	Tuning ratio	Channel coverage, Mc	Power output, watts
3C22.....	Lighthouse triode	External	To 1500	1500	60-10
2C39.....	Lighthouse triode	External	To 2500	2500	50-10
5J30.....	Magnetron	External	350-750	2.15:1	400	150
6J21.....	Magnetron	Internal	2400-3600	1.5:1	1200	1000
4J60.....	Magnetron	Internal	2700-3000	1.11:1	300	60
368AS.....	Triode	External	To 600	600	10-3
Resnatron...	Tetrode	Internal	400-600	1.5:1	200	50,000

13-3. Modulation.—In order to establish a maximum of interference on a radar channel, it is necessary to generate a modulated carrier wave having wide-band components of substantial magnitude at frequencies several megacycles removed from the carrier frequency. Experiments showed that to meet these requirements the most effective, readily generated signal consisted of a carrier wave, heavily amplitude modulated by random noise. The modulator output preferably consists of a relatively uniform spectrum of video noise extending from low frequencies to high frequencies of the order of 2 to 5 Mc. It was experimentally determined that linearity of amplitude response was not essential, and consequently this point receives little attention in the material that follows. On the other hand, a uniform frequency response was always desirable, and the discussions of the various factors affecting band width of the transmitted signal are given in some detail. In many cases the amplitude modulation was accompanied by unintentional incidental frequency modulation of occasionally as much as several megacycles. It was found, however, that from the point of view of effectiveness of the signal, the incidental frequency modulation was for the most part harmless and in some cases beneficial, so that little or no work was undertaken to provide for stability of carrier in the presence of modulation. Nevertheless, as a part of the general investigation to determine the most effective type of modulation, several frequency-modulation studies were undertaken, and the technical results of this series of investigations are rather completely covered in Secs. 17-15 to 17-26.

13-4. Transmitting Equipment Types and Requirements.—Because of the fact that the equipments developed were to be produced and employed in quantity, military requirements were established relating to their ease of operation and maintenance. It was also desirable to keep

the number of equipment types to a minimum and to provide for a maximum of frequency coverage in any one type. These considerations resulted in the general employment of modulated-oscillator systems having a minimum of components, *viz.*, a power supply, a modulator, and an oscillator. Some work was done on master-oscillator power-amplifier combinations and some on single-side-band systems, but these schemes generally were restricted to operating frequencies below 200 Mc. The results of this particular work represent no particular contributions to the art and are as a result not reported here. This is not to imply that power-amplifier systems employing triodes and tetrodes are not important or practical at frequencies up to 2000 to 3000 Mc, but rather that the oscillator treatment given is of a sufficiently general nature to permit interpretation in terms of amplifier and frequency-multiplier performance. Emphasis on the modulated power oscillator is additionally clear when it is remembered that a substantial portion of the power and frequency range was covered through the employment of the appropriate magnetrons, which are inherently power oscillators.

13-5. Summary of Progress in the Field of Power Generation.—The completed developments include the following types of equipment:

1. *Triode Systems.*—Power oscillators of 10 to 50 watts output, depending on frequency, operating in coaxial external cavities tunable from a low-frequency limit dependent only upon tolerable cavity length, to a high-frequency limit of 1500 Mc in the case of the 3C22, or 2500 Mc in the case of the 2C39, and with plate efficiencies of 60 per cent or more at the low-frequency end of the range and 15 to 20 per cent at the high-frequency limit.

2. *Tetrode Systems.*—A tetrode power oscillator of 50,000 watts output unmodulated, tunable internally from 400 to 600 Mc at plate efficiencies of 60 to 70 per cent.

3. *Magnetrons I.*—A series of three magnetrons operable between the frequencies of 90 and 1200 Mc, externally tuned, and having power outputs of 150 watts at plate efficiencies of 30 to 40 per cent.

4. *Magnetrons II.*—Two magnetrons of 1000 watts output: the first, an externally tuned tube operable from 90 to 250 Mc, with a conversion efficiency of 40 to 50 per cent, and the second, an internally tuned tube covering the range 2400 to 3600 Mc with a plate efficiency of 50 to 60 per cent.

5. *Magnetrons III.*—A series of six internally tuned magnetrons covering the frequency range 2300 to 4000 Mc with power outputs of 50 watts at 40 to 50 per cent efficiency.

In addition to these completed developments there are under development a number of magnetrons of various sizes and frequency ranges. In the case of these magnetron developments most of the basic work is complete; problems remaining are to a large extent the elimination of

parasitic resonances in the cavities, which adversely affect operation in "spots" in the tuning range of the tube.

Triode and tetrode developments are also under way to cover more adequately the power gap between the resnatron and disk-seal triodes of the 3C22 type. Activity in the new television bands will, no doubt, accelerate these developments and result in satisfactory triodes and tetrodes having 1 to 5 kw of power output in the frequency ranges below 1000 Mc.

The following chapters have been written to establish a better understanding of the principles involved in the generation, modulation, and measurement of substantial amounts of high-frequency power. It is well to remember that the work was carried on under the extreme pressure of a wartime emergency and represents the results of the most direct approach to the solution of the problem. Inevitably, topics of great interest had to be neglected in the interest of speed, and in these cases the gaps in the treatment of the problem will be obvious. This, of course, is justifiable, and will indicate to the reader the number and types of problems that will be subject to continued research and development.

CHAPTER 14

TRIODE AND PENTODE ULTRAHIGH-FREQUENCY OSCILLATORS

By E. A. YUNKER, J. F. BYRNE, J. G. STEPHENSON, AND M. B. ADAMS

14-1. Introduction.—When the methods and calculations applicable to triode and pentode tubes at low and medium frequencies are used in the v-h-f and u-h-f region for the design of oscillators and amplifiers, various difficulties are encountered. In general, the power obtained does not agree with calculated values, and considerable experimentation is usually necessary before optimum results are obtained from a given tube and circuit configuration. One factor contributing to this situation is electron transit time within the tube. Another is that, when the frequency is high, stray, or not obvious, inductances and capacitances have marked effects on the oscillation frequency as well as in causing frequency jump-

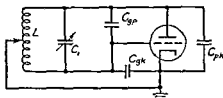


FIG. 14-1.—A simple low-frequency Hartley oscillator

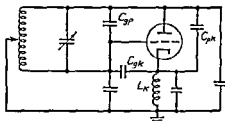


FIG. 14-2 —Modified Hartley oscillator.

ing and holes, or regions in the frequency spectrum in which the power output is zero or quite low. In general, the optimum proportions of the circuit are not always readily calculable, and considerable use of trial and error may be necessary.

The circuit of a simple Hartley oscillator, as used at low frequencies, is shown in Fig. 14-1. For the sake of simplicity, power and bias sources are not shown. Here the tuning capacitance C_1 is large compared with the interelectrode capacitances of the tube, so that the latter may be neglected. To cause this circuit to oscillate, it is necessary to feed back a voltage of proper phase and magnitude from the plate circuit to the grid. The exact conditions for oscillation are given in a number of textbooks and will not be taken up here. At very high frequencies, however, the inductance of the leads in the tube becomes important, with the result

that the cathode of the circuit of Fig. 14-1 is not at ground potential. This and the fact that the tube capacitances are then important causes a change in the circuit, resulting in that of Fig. 14-2. The circuit is then no longer that of a Hartley oscillator. The voltage on the grid now depends not only on the position of the cathode tap on the coil, but also on the several interelectrode and ground capacitances and on the cathode inductance. It is to be noted that while the grid and plate leads also have inductances they are parts of the tuned circuits and so do not require special consideration. As may be expected, in view of the complexity of the circuit and the relative magnitudes of the stray or incidental circuit parameters, the performance will differ considerably from that of the circuit of Fig. 14-1.

The difficulties thus far discussed are caused principally by the length and, therefore, inductance of the cathode leads. It is not possible to

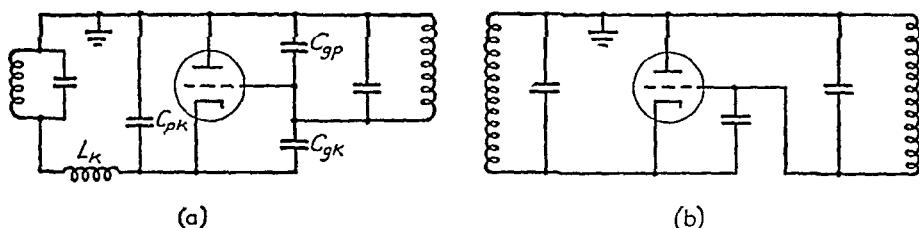


FIG. 14-3.—Grounded-plate oscillator circuit.

make these leads as short as might be desired for electrical reasons, because the temperature difference between cathode and tube envelope dictates a minimum length. In general, any point on an oscillatory system or other network can be grounded, the only effect of a change of ground point being a difference in the distribution of voltages with respect to ground and the effect of stray parameters. By properly choosing the ground point, troublesome stray capacitances, for example, may be in effect removed from the circuit. Thus, in the circuit of Fig. 14-2, four of the capacitances shown terminate on the plate of the tube or the connection to it. By grounding the plate and connecting resonant circuits between plate and grid and between plate and cathode, the necessarily long cathode leads are made part of the tuned plate-cathode circuit, and several of the capacitances are in parallel with tuning capacitances. The result is the circuit of Fig. 14-3a, which reduces to that of Fig. 14-3b when the filament-lead inductance L_k and the plate-cathode capacitance C_{pk} are added to the plate-cathode tuned circuit and the grid-plate capacitance C_{gp} to the grid-plate tuned circuit. When coaxial or other resonant-line circuits are used, as they probably would be for those frequencies for which the inductance of the cathode lead is objectionable, the circuit becomes that of Fig. 14-4. By proper design of the tube the inductance

of the plate lead may be made extremely small. The principle of selecting a circuit ground point such as to minimize, for a given tube, the effect

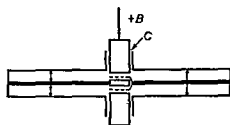


FIG. 14-4.—Coaxial-line equivalent of the oscillator of Fig. 14-3.

of lead inductances and capacitances gives rise to a number of circuit types such as grounded grid with tuned anode and cathode, and grounded plate with tuned grid and cathode. Indeed, there is in general a given type of circuit that will work best with each type of tube, and conversely, the design of a particular tube should be based on some particular type of circuit.

14-2. Oscillator Analysis.—A further fundamental difference exists between the Hartley circuit of Fig. 14-1 and the circuits of Figs. 14-3 and 14-4, as shown by the following considerations.

1. In a triode oscillator or amplifier, efficient power conversion occurs when the plate-current pulses flow at or near plate-voltage minimums. With low values of transit time, this means that grid-voltage maximums and plate-voltage minimums occur simultaneously, *i.e.*, the alternating plate and grid voltages are 180 deg out of phase.

2. In oscillator circuits reducible to a Δ - (or π -) network, where only the resultant grid-plate, grid-cathode, and plate-cathode impedances need be considered, the combination of these impedances, acting as a π -section, must be such as to produce approximately a 180-deg phase shift through the network.

3. Two possible π -networks are shown in Fig. 14-5, each of which, with appropriate numerical values attached to the reactances, will produce the 180-deg phase relation between plate-cathode and grid-cathode voltages. The simplified oscillation conditions are that $X_1 + X_2 + X_3 = 0$ and $X_1 = X_3/\mu$. Clearly, X_1 and X_3 must be of the same sign and X_2 must be opposite in sign to X_1 and X_3 and equal to their sum. In these two π -networks, which represent all the possibilities as far as oscillation is concerned, the reactances are the total equivalent grid-plate, plate-cathode, and grid-cathode reactances, including strays, connected reactances, and interelectrode capacitances. Admittedly, the effect of lead inductance (a common impedance effect) has been neglected so far, but without loss of generality, as will be shown later.

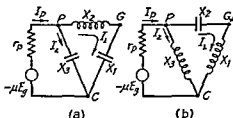


FIG. 14-5.—Oscillator classes: (a) Class I, (b) Class II.

An advantage of the Class I oscillator is apparent when the effects of transit time and circuit losses (including useful power output) are considered. To illustrate this point it is convenient to construct vector dia-

grams of the two classes of oscillators. Figures 14-6a and b show these completed vector diagrams. They are constructed by taking the plate-cathode voltage E_p as reference vector. In the Class I oscillator diagram of Fig. 14-6a, the current I_1 , of Fig. 14-5a, flowing through the grid-plate and grid-cathode impedances is obviously not quite 90 deg behind the plate-cathode voltage because of power absorbed in this circuit branch. It lags the voltage E_p because the absolute value of the equivalent grid-plate reactance must be greater than that of the grid-cathode reactance. Since grid driving power is ordinarily required, the grid-cathode voltage E_g similarly lags this current by not quite 90 deg. Accordingly, E_g lags the voltage E_p by less than 180 deg. The equivalent voltage $-\mu E_g$ is drawn in opposite direction from E_g and is shown dotted in the diagram. It is apparent from Fig. 14-6 that the vector $-\mu E_g$ is equal to the sum of the vector E_p plus the vector $I_p r_p$. A remaining circuit

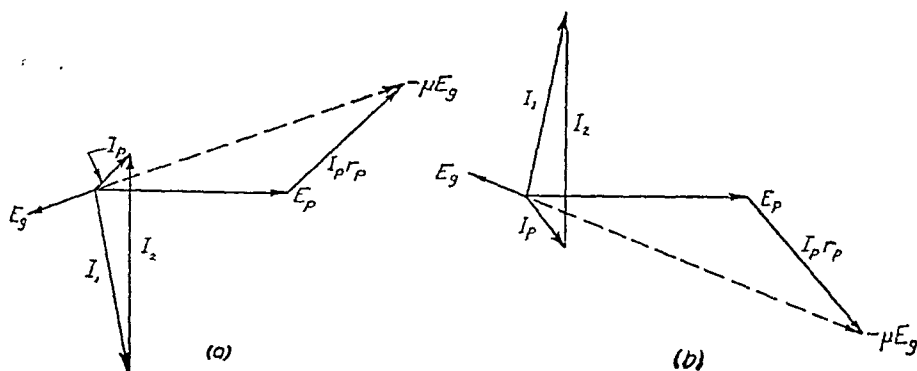


FIG. 14-6.—Vector diagrams: (a) Class I oscillator; (b) Class II oscillator.

condition from Fig. 14-5 that must be satisfied is that the plate current $I_p = I_1 + I_2$. This condition is established on the vector diagram by drawing a vector leading E_p by 90 deg, originating at the end of the current vector I_1 and continuing until it intersects a line drawn from the origin parallel to the vector $I_p r_p$. This vector is the current I_2 , and the sum of $I_1 + I_2$ is the plate-current vector I_p . It should be noted that under these conditions the plate current I_p is definitely not in phase with the plate voltage E_p and cannot be in phase with it as long as there is load in the grid circuit. However, the higher the Q of the circuits, the more nearly in phase these quantities become.

A similar diagram for the Class II oscillator is illustrated in Fig. 14-6b. It is noticed that in the Class I oscillator the voltage $-\mu E_g$ in general leads the plate-cathode voltage E_p , while in the Class II oscillator it lags E_p . When the effects of transit time are considered, the vector $-\mu E_g$ is conveniently drawn with more than 180-deg lag relative to E_g , the additional amount being equal to the transit angle. From Fig. 14-6a it is

clear that the vector $-\mu E_g$ will be rotated clockwise. To a certain point this effect can be beneficial. In the Class II oscillator, however, any clockwise rotation of $-\mu E_g$ will result in greater phase differences between plate current and plate-cathode voltage with a consequent reduction in oscillator efficiency.

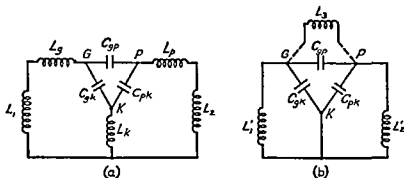


FIG. 14-7.—Effect of lead inductance in tuned-grid tuned-plate oscillator (a) actual circuit; (b) equivalent π -section.

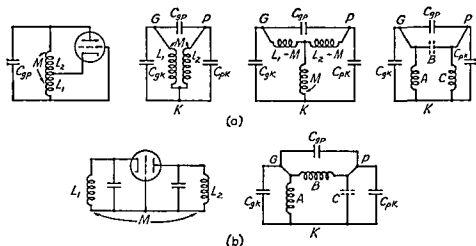


FIG. 14-8—Oscillators with inductive coupling: (a) Hartley, $A = \frac{L_1 L_2 - M^2}{L_1 - M}$, $B = \frac{L_1 L_2 - M^2}{L_1 - M}$, $C = \frac{L_1 L_2 - M^2}{L_1 - M}$; (b) grounded-grid, $A = \frac{L_1 L_2 + M^2}{L_1 \pm M}$, $B = \frac{L_1 L_2 + M^2}{L_1 \pm M}$, $C = \pm \frac{L_1 L_2 + M^2}{M}$.

From the foregoing discussion it is apparent that Class I oscillators are superior to Class II oscillators at ultrahigh frequencies since electron transit angle, if not too large, may be beneficial to Class I oscillator performance, whereas it is always detrimental to Class II oscillator performance.

It has been stated previously that the general considerations pertain-

ing to Class I- and Class II-oscillator operation are valid when the effects of lead inductances are taken into account. Figure 14-7 is a simplified diagram of a tuned-grid tuned-plate oscillator in which the interelectrode capacitances and lead inductances are shown. It is significant to note in this case that the inductance of the cathode lead is common to the grid-cathode and the plate-cathode circuit, while the inductance of the grid and plate leads can conveniently be thought of as a part of the external circuit. Inspection of the diagram shows that the cathode inductance forms the shunt arm of a T-section consisting of $L_g + L_1$ and $L_p + L_2$ as series arms. At any particular frequency, say the operating frequency, this T-section may be replaced by its equivalent π -section, consisting of the two shunt inductances L_1' and L_2' and the series inductance L_3 . In this way the circuit may be reduced to the simplified circuit of Fig. 14-5b. It will be noticed that in this case the equivalent π -section shows the inductance L_3 in parallel with the grid-plate capacitance, and it might be assumed that at some frequency the equivalent parallel reactance of this combination would reduce the feedback to a point where the system would fail to oscillate. This is, indeed, the case, and the following general statement can be made: the effect of lead inductance to the common electrode (not necessarily the cathode) is such as to reduce the feedback by increasing amounts as the operating frequency of the oscillator is increased. This statement is true regardless of whether the oscillator is operating as a Class I or Class II oscillator.

It can be shown, by means of the series of transformations indicated by the circuits of Fig. 14-8, that oscillators using mutual-inductance feedback can be classified as either Class I or Class II oscillators. Thus in Fig. 14-8a the Hartley circuit is seen to be of the Class II group, while the grounded-grid circuit of Fig. 14-8b, which is equivalent to a grid-separation coaxial-line oscillator such as that shown in Fig. 15-15, is a Class I oscillator. Several other standard circuits are analyzed in a similar manner in Fig. 14-9.

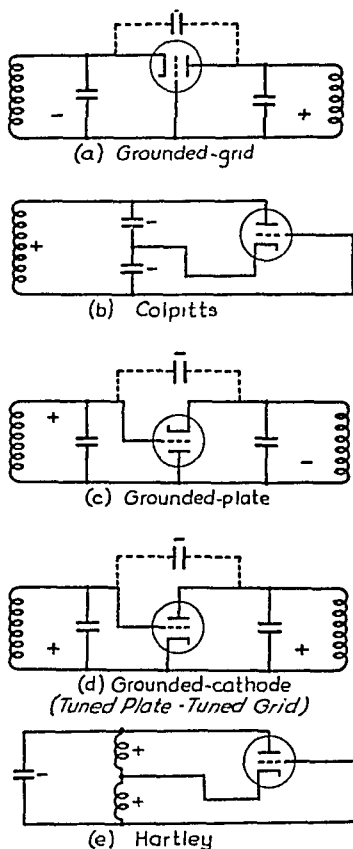


FIG. 14-9.—Class I and Class II oscillators. A plus sign indicates that an inductive reactance is required for oscillation; a minus sign, a capacitive reactance. (a), (b), and (c) are Class I; (d) and (e), Class II.

14-3. Choice of Tubes and Circuits.—The most important single factor in selecting the proper circuit for use with a given tube, for operation at frequencies approaching the upper frequency limit of the tube, is the configuration of the tube itself. As discussed in Sec. 14-1, any one

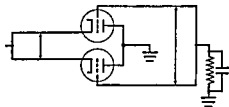


FIG. 14-10.—Grounded-plate parallel-line push-pull oscillator.



FIG. 14-11.—Grounded-plate parallel-line angle-end oscillator.

of the three tube elements may be selected as the common electrode. Since the inductance of the leads to the common electrode should be kept as low as possible, a tube such as shown in Fig. 15-2*d* should be used in common plate-grounded plate circuits, such as shown in Figs 14-4,

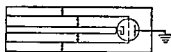


FIG. 14-12.—Grid-separation coaxial-line oscillator.

14-10, and 14-11. The grid-plate and plate-cathode circuits should be coaxial lines for the highest frequencies. They may, however, be parallel lines since a circuit using two tubes in push-pull as shown in Fig. 14-10, with the plate electrodes strapped together and grounded, will work well with parallel lines to relatively high frequencies. The tube of Fig. 15-2*j* was designed for common-plate operation. A common-grid coaxial-line oscillator is best for the tubes of Figs 15-2*b* and *f*. Because the plate electrode must be air-cooled, it is ordinarily connected to the outermost

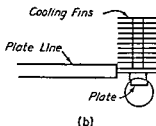
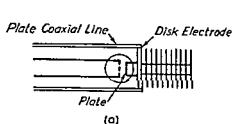


FIG 14-13.—Effect of electrodes and plate cooling fins of disk-electrode type of tube when used with (a) coaxial-line oscillator; (b) parallel-line oscillator.

tube of the coaxial system (Figs. 14-12 and 15-15), which for reasons of safety is usually grounded. The tube of Fig. 15-2*a* is generally used with the innermost tube of a coaxial system connected to the plate electrode. None of the tubes of Fig. 15-2 lend themselves well to parallel-line sys-

tems because the large disk electrodes, designed for coaxial systems to reduce inductance to a minimum, become lumps of capacitance when connected to the ends of parallel lines. Although the disk electrode, used in a coaxial-line oscillator, forms part of the line, as shown in Fig. 14-13a, the same electrode in a conventional parallel-line circuit, Fig. 14-13b, becomes an enlargement to the line and thus adds to the capacitance between lines and to ground. This is particularly true for plate electrodes that have fins for the dissipation of heat. It is possible to devise parallel lines into which the disk-electrode tubes may fit, but these are in general cumbersome because of the large diameter of line required. On the other hand, tubes such as WE 368A, 368AS, and 388A are not well suited for coaxial-line circuits, nor are most of the conventional transmitting tubes.

In u-h-f oscillators, electrical characteristics are closely identified with and dependent upon mechanical features. For this reason, further discussion of oscillators will be taken up under two headings: oscillators and amplifiers using parallel lines, and those employing coaxial lines. The latter class is discussed in Chap. 15.

14-4. Parallel-line Oscillators.—Parallel lines as resonant circuits have a number of advantages over other types of resonators for use in v-h-f oscillators. They are relatively simple and cheap to construct, their electrical length may be readily changed by short-circuiting bars, and when they are used with appropriate types of tubes, the connections between lines and tube terminals may be short and direct. Furthermore, these connections and the portions of the tube leads inside the envelope become parts of the resonant-line system. For very high frequencies the tube leads may constitute the principal part of this system but are largely inaccessible for purposes of coupling out power. In some cases the portion of the circuit from which power is to be coupled may be operated at a multiple length of the shortest possible length of line, *e.g.*, at $3\lambda/4$ rather than $\lambda/4$. However, this usually cannot be done in parallel-line oscillators because the frequencies involved are likely to be too high for effective use of this type of circuit.

Parallel lines are particularly suited for push-pull circuits, such as shown in Figs. 14-10, 14-14, 14-16, 14-17, 14-21, and 14-23. In these circuits plate and bias voltages are applied at low r-f potential points. Furthermore, the resonant circuits are isolated from each other with respect to d-c voltages so that parallel feed, which for wide tuning ranges often introduces problems with chokes, is not necessary.

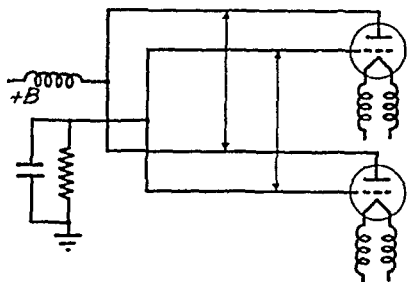


FIG. 14-14.—Parallel-line, push-pull oscillator using cathode chokes.

Some control may be exercised over the Q and characteristic impedance Z_0 (Secs. 15-7, 15-11, and 15-13) of the parallel-line oscillator. However, these factors are often subordinated to mechanical considerations. Thus in an oscillator using the circuit of Fig. 14-14 with a pair of 368AS tubes, the size of tubes, spacing of pins on the tubes, and frequency range required are such as to dictate line spacing between about 1 and 2 in. and line diameters between $\frac{3}{16}$ and $\frac{1}{2}$ in. Thus the possible values of the characteristic impedance Z_0 of the line range between 166 and 660 ohms.

The frequency range of the parallel-line oscillator is limited on the upper end by the length of the tube leads and the capacitances of the tube, or by electron-transit-time losses at the permissible voltages used on the tube. Which is the more important depends, of course, on the tube design. With u-h-f tubes such as the WE 368AS, the line limitation usually predominates; for the Eimac 35TG, for example, it is transit time. The low-frequency limit for parallel lines is usually determined by the rather inconvenient lengths of lines required for low-frequency operation, although this objection may be removed by coiling or bending the lines. In the latter case, however, the mechanical problems of tuning may become troublesome. As the frequency at which the oscillator is to be operated is reduced, the advantage of parallel lines over circuits using coils and variable capacitors or variable inductances with distributed capacitance (see Sec. 14-8) becomes less.

Power may be readily coupled out of parallel-line oscillators by means of devices such as shown in Figs. 16-5 and 16-6 and discussed in Sec. 16-4 and 16-5. In general, the coupling device, for use on push-pull circuits, should be of the balanced type. This balance includes the maintenance of electrostatic symmetry to the resonant lines of the oscillator.

14-5. 450- to 720-Mc 8-watt Oscillator.—The mechanical parts of an oscillator using the circuit of Fig. 14-14 are shown in Fig. 14-15. Two WE 368AS "doorknob" tubes are employed. Feed screws between lines, shown in Fig. 14-15, are used to move the shorting bars. These may be moved simultaneously by gearing the two screws together, although some form of mechanical compensation is usually required for tracking over a wide range because the two shorting bars do not in general move at the same rate with respect to frequency. The frequency range of the unit shown is 450 to 720 Mc; the power is about 8 watts at the low-frequency end of the tuning range and 4 watts at the upper. The resonant lines and feed screws are supported by a single block of mica-filled bakelite at the ends opposite the tubes. The filament chokes are formed by winding two wires, one for each filament lead, in parallel around a cylindrical form,

The series-stub-tuned loop discussed in Sec. 16-4 is a good device for coupling power from a parallel-line oscillator. This loop may be tuned independently or may be ganged with the oscillator tuning control. For independent tuning the loop assembly should be movable along the resonant lines so that the loop may be placed in the proper position for the required degree of coupling. This position is ordinarily near the short-circuiting bar. If the coupler tuning is to be ganged with the oscillator controls, the coupling loop and tuning stub may be carried along with the plate-line short-circuiting bar so that the loop is always above the latter. Tuning of the stub is then accomplished by attaching the stub

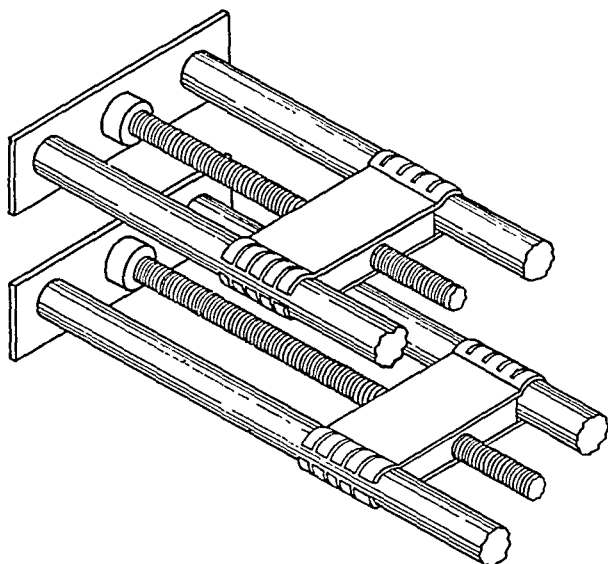


Fig. 14-15.—Tuning mechanism for oscillators of Figs. 14-14 and 14-17.

tuning piston to the stationary parts of the oscillator through a slot in the outer coaxial line of the stub.

The inductively coupled balun, shown in Fig. 16-6 and discussed in Sec. 16-5, is also a good device for use with the oscillator of Figs. 14-14 and 14-15. It is less critical in tuning than the tuned loop and may be easily driven from the plate-line short-circuiting bar for single-dial operation because the rate of change of position of the balun short-circuiting bar, with respect to frequency, is nearly the same as that of the plate-line short-circuiting bar.

An oscillator of the type discussed above will not in general operate with either the plate or the grid lines on a $3\lambda/4$ mode because the two circuits are coupled by the proximity of the plate and grid lines and because circuit losses are higher at higher frequencies. Although coupling due to proximity of the plate and grid circuits is, of course, not essential for

operation of this type of oscillator, this arrangement is used in the oscillator of Figs. 14-14 and 14-15 in order to obtain compactness and simplicity of tuning. As a result, if one line is lengthened as for $3\lambda/4$ mode operation at a certain frequency, oscillation at a lower frequency corresponding to the $\lambda/4$ mode usually results.

14-6. 250-to-550-Mc 20-watt Oscillator.—The circuit of Fig. 14-16 is similar to that of Fig. 14-14 except that double-lead tubes are used and that a tuned circuit is connected to each of the leads. Thus the tube is located at the middle of a pair of lines electrically $\frac{1}{2}$ wavelength long when loaded in the middle by the tube capacitances. This arrangement divides the total tube capacitance between two circuits, *i.e.*, one to the right and one to the left of the tube in Fig. 14-16, thus permitting a higher frequency at the limiting minimum length of line. Such a device has, however, the disadvantage of increasing the complexity of the tuning

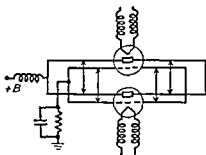


FIG. 14-16 — Double-resonant-circuit oscillator for reducing the effect of interelement capacitance of tubes.

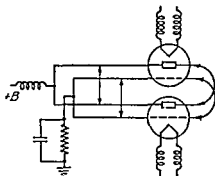


FIG. 14-17 — Use of back-lines for reducing the effect of tube capacitances.

system. An alternate method is to substitute lines of fixed length for the variable-length lines on the right-hand side of Fig. 14-16, as shown in Fig. 14-17. Spring finger contacts are used on the ends of the *back-lines*, which are made of spring brass strips, so that they may be clipped to the tube pins as required. For a wide-range oscillator these lines should be $\frac{1}{4}$ wavelength long, electrically, at the highest frequency of operation, when loaded with half the tube capacitance. The fact that the tube elements will then be at the middle of the half-wave system at this frequency favors the high-frequency end of the range where the power is inherently lowest. As the frequency is lowered, the electrical middle, or voltage antinode, will shift away from the tube elements. This fact determines the lower limit at which the lines are useful.¹

The frequency range of an oscillator that uses a pair of WE 388A tubes, in the circuit of Fig. 14-17, is 250 to 550 Mc with a power output of about 25 watts at the lower frequency limit and 8 watts at the upper.

¹ See Sec. 23-2 for additional discussion of this action.

The modulation band width is 7 Mc. The back-lines are used above 450 Mc.

By loading the plate and grid lines by means of small coils, shown in Fig. 14-18, connected between the ends of the lines and the tube terminals in place of the usual strap or connector, the frequency of a parallel-line

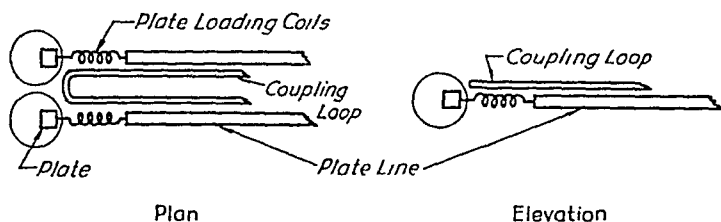


Fig. 14-18.—Loading coils for reducing the frequency of parallel-line oscillators.

oscillator may be reduced considerably. If a loop coupling system is used with the oscillator of Fig. 14-18, the axis of the coils should be parallel to the plate lines and should be an extension of the lines, so that the loop may pass between. This is so because considerable magnetic coupling is obtained from the magnetic field of the axial component of current in the coil.

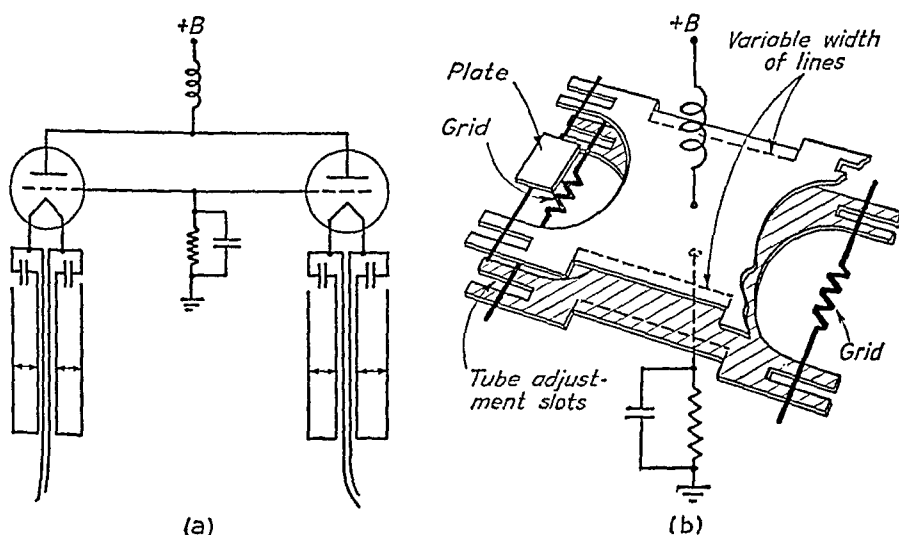


Fig. 14-19.—Push-pull oscillator using parallel plate-type lines.

14-7. Oscillator Using Lines of Variable Characteristic Impedance.—

The circuit shown in Fig. 14-19 is particularly suited to the RCA 8012 and similar narrow cylindrical tubes with double leads. The resonant circuit consists of parallel plate-type lines, with faces adjacent, with a tube at each end. Since the equivalent reactance of a resonant parallel

line, open at the receiving end, is equal to $Z_0 \cot (2\pi l/\lambda)$, where Z_0 is the characteristic impedance of the line and l , its length, the resonant frequency of such a line loaded at the end by tube capacitance may be varied by varying Z_0 . This is done in the oscillator of Fig. 14-19 by varying the width of the line, keeping the length and spacing fixed. Thus, when the line is narrow, the frequency is low, and vice versa. This arrangement does not require that the tubes be moved during the tuning process. For additional frequency range, the tubes may be moved longitudinally along the line in slots, as shown in Fig. 14-19b.

A very satisfactory coupling system for this oscillator consists of a loop placed above the anode line of the oscillator so as to intercept the magnetic field around that line. It may be swung around an axis parallel to and above the plate line, for adjustment of the degree of coupling.

Tunable coaxial cathode lines are used in this oscillator. These are operated from the main tuning control through a system of gears so that their length, at each frequency, is such as to give the best operation. A small mica capacitor is used between each filament lead and the center conductor of the coaxial line in order to maintain all parts of the filament at nearly the same r-f potential.

14-8. Coiled Lines.—For frequencies below, say, 100 Mc the length of line required for a parallel-line oscillator becomes objectionably large when compactness is required. Compactness can be achieved, however, by winding the line into a coil or spiral. Such an arrangement, compared with a lumped-parameter circuit, has the advantages of wide tuning range and low operating Q , the latter being important in applications involving broad-band modulation. The mechanism for short-circuiting out the unused portion of the line or coil is, however, usually more complicated than that required for simple parallel lines.

It may be argued that a line wound into a spiral is an inductance and that a circuit using it is a lumped-parameter circuit. However, in the absence of a relatively large external lumped capacitance, the circuit exhibits the characteristics of one having distributed inductance and capacitance. The coiled lines are, of course, coils in a physical sense and for convenience are so designated in the following sections. Figures 14-20 to 14-26 illustrate several methods of constructing variable coiled lines and also methods of coupling them to a load.

14-9. Spiral-line Oscillator.—The oscillator shown in Fig. 14-20, using a pair of 3C24 tubes, operates over the frequency range 60 to 340 Mc with a power output of 50 watts at the lower frequency and 20 watts at the higher. A tuned-plate circuit with capacitance feedback is used, as shown in Fig. 14-21. The capacitances of the cylindrical capacitors, which are fabricated from polystyrene cylinders and brass tubes, are chosen so as to give sufficient feedback at the lowest frequency. Thus

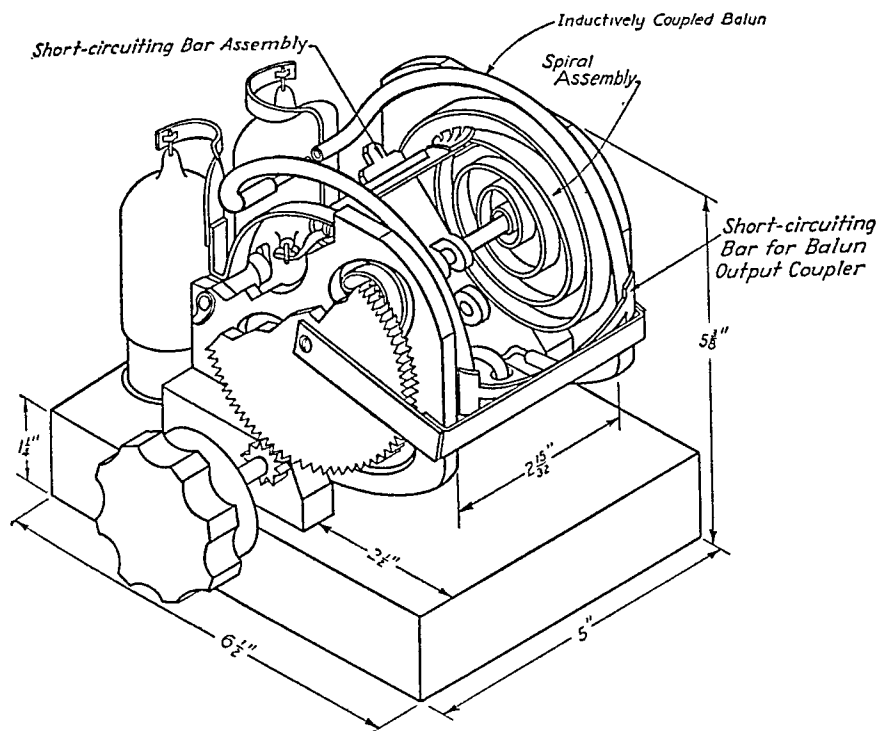


FIG. 14-20.—Spiral-line oscillator with inductively coupled balun output.

the feedback is increased with frequency. On the oscillator shown, this capacitance was $7 \mu\text{f}$.

Tuning is accomplished by a short-circuiting bar between corresponding points on the two spiral lines. Radial movement of this bar is along an arm attached to a shaft at the center of the spirals. Spring fingers serve to make contact between the short-circuiting bar and the flat metal strips of the spirals.

The output system of the oscillator of Fig. 14-20 is an inductively coupled balun (see Sec. 16-5), bent so as to parallel the plate spiral for the half turn nearest the tube. Because of the broad frequency response of the balun coupling system, this length is sufficient to give satisfactory coupling over the whole range of the oscillator. The balun is tuned by a short-circuiting bar similar to

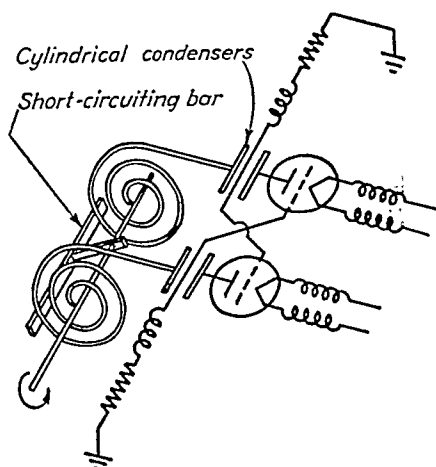


FIG. 14-21.—Schematic diagram of oscillator of Fig. 14-20.

of the oscillator. The balun is tuned by a short-circuiting bar similar to

that used on the plate spiral. The two tuning controls may be ganged through a suitable gear reduction between the multiturn plate spiral and the half turn of the balun.

14-10. Helical-line Oscillators.—A *rollo-type* coil or line, in which the center portion is short-circuited out by rolling contact wheels, is shown in Fig. 14-22. The wheels are mounted so that they can slide freely along a supporting rod on which they are pinioned and so that pressure contact can be made to the coil wire. Because of the opposite winding direc-

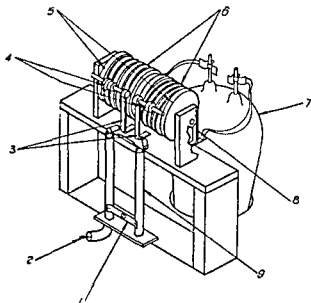


FIG. 14-22.—Rollo coil variable-inductance tank circuit: (1) balun short-circuiting bar (adjustable); (2) 50-ohm line to load, (3) spring supporting arms; (4) rollo contacts with grooved rim which fits over wire of coil. Contacts free to rotate about pinion arm, and to slide axially, (5) slip rings (brush contacts on opposite side of coil); (6) rollo coil (rotates) (opposite winding direction in two halves of coil); (7) amplifier tube; (8) spring brush contact to slip ring; (9) capacitively coupled balun output coupler. Note: All coils, slip rings, and contacts silver-plated.

tion in the two halves of the coil, the rollo contacts move symmetrically with respect to the center of the coil as it is rotated. Slip-ring contacts are made to the ends and center of the coil. The winding consists of silver-plated wire on an MF formica form. All slip rings and contacts are silver-plated. This coil, with the associated stray and interelectrode capacitances, has a tuning range of 90 to 220 Mc when used with a type-832 tube. A balun output transformer is capacitively coupled to the center point of the tank coil (see Sec. 16-5).

Figure 14-23 shows another form of variable inductance or spiral line in which contact to the rotating coil is made with small sliding fingers that nest between two adjacent turns on the coil. The sliders themselves

are free to move symmetrically with respect to the center of the coil as it is rotated. A tuning range of 25 to 105 Mc was obtained with the coil of Fig. 14-23, when used with a type-829B tube. Output coupling is accomplished by means of a small r-f transformer, the primary of which consists of a few turns of heavy wire in series with the tank inductance at the center point. This amounts to close loop coupling to the tank circuit.

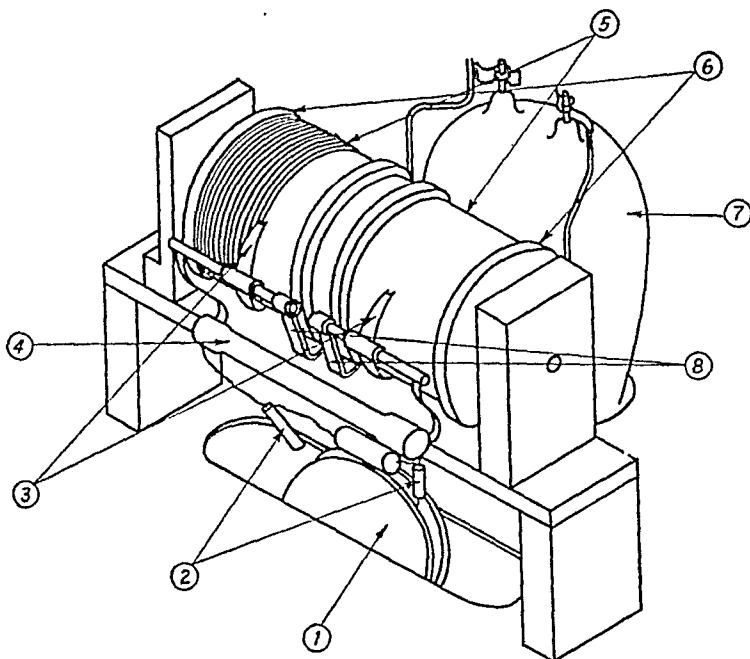


FIG. 14-23.—Variable-inductance tank circuit with spring finger contact: (1) r-f output transformer; (2) capacitors for isolating antenna circuit from d-c plate voltage; (3) spring contacts nesting between two adjacent turns of coil (not rotatable but free to slide axially); (4) r-f choke for plate voltage feed; (5) rolo coil (rotates) (opposite winding direction in two halves of coil); (6) slip rings (brush contacts not visible); (7) amplifier tube 829B; (8) spring material. Note: All coils, slip rings, and contacts silver-plated.

A variant of the coil of Fig. 14-23 is shown in Fig. 14-24. The inductance is made of edge-wound copper strip. The spring short-circuiting contacts slide along the outer edge as they rotate around the coil. This coil was designed for an amplifier tank circuit and, with an HK257B tube, tunes from 25 to 100 Mc.

Figure 14-25 illustrates a further modification of the coil shown in Fig. 14-23. A short-circuiting sleeve cuts out end turns of the inductance as the coil is rotated. This method is mechanically somewhat less satisfactory than the others mentioned, but in the oscillator shown will give a frequency range of 70 to 270 Mc when used with a type-829B tube. The output coupler consists of a balun (see Sec. 16-7) conductively coupled

to the inner ends of the two halves of the coil. The balun is bent in a curve such that the change in frequency is linear with the angle of rotation of the arm carrying the balun short-circuiting bar.

A type of rollo coil in which the rolling contacts are spring-mounted on hairpins to press against the inside of the coil is shown in Fig. 14-26. The entire hairpin and roller assembly rotates within the fixed coils, short-circuiting out end turns. The coils are of heavy and rigid construction, and all parts are silver-plated. This coil and its associated

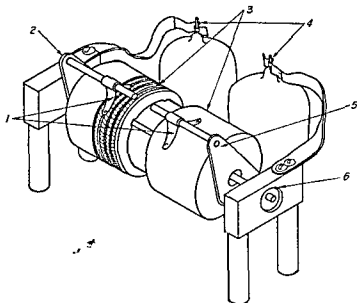


FIG. 14-24.—Edge-wound ribbon variable-inductance tank circuit: (1) spring short-circuiting contacts riding on edge-wound ribbon coils and free to slide axially along coil as well as to rotate completely around it; (2) rotating arm; (3) edge-wound inductance coil (fixed) (opposite winding direction in two halves); (4) plate leads of amplifier tubes; (5) rotating arm driving spring contact around fixed coil; (6) spring contact maintained between rotating arm and stationary bushing. Note: All coils, slip rings, and contacts silver-plated.

amplifier circuit, using a type-HK257B tube, have a tuning range of 85 to 150 Mc. A similar coil having a smaller number of turns tunes from 140 to 210 Mc with Eimac 35TG tubes. Output coupling is accomplished by arranging an inductively coupled balun transformer so that it lies in close proximity to a loop formed in series with the tank inductance at its center, as shown in Fig. 14-26. Balun couplers are discussed in greater detail in Sec. 16-5.

14-11. Notes on Design and Construction.—Since open parallel lines radiate electromagnetic energy when excited, it is necessary to shield these lines in an oscillator. This is particularly true for very high frequencies because the spacing between lines, which is dictated by mechanical and impedance considerations, becomes an appreciable fraction of a

wavelength at these frequencies. Radiation is also severe from oscillators of the type shown in Fig. 14-19 because the resonant lines are, when end loaded by the tubes, electrically $\frac{1}{2}$ wavelength long and thus act as dipole radiators. Furthermore, the broad plate line shields the grid line below it so that the cancellation of radiation due to the phase opposition of the two lines is reduced in the upward direction. This oscillator works very poorly when unshielded.

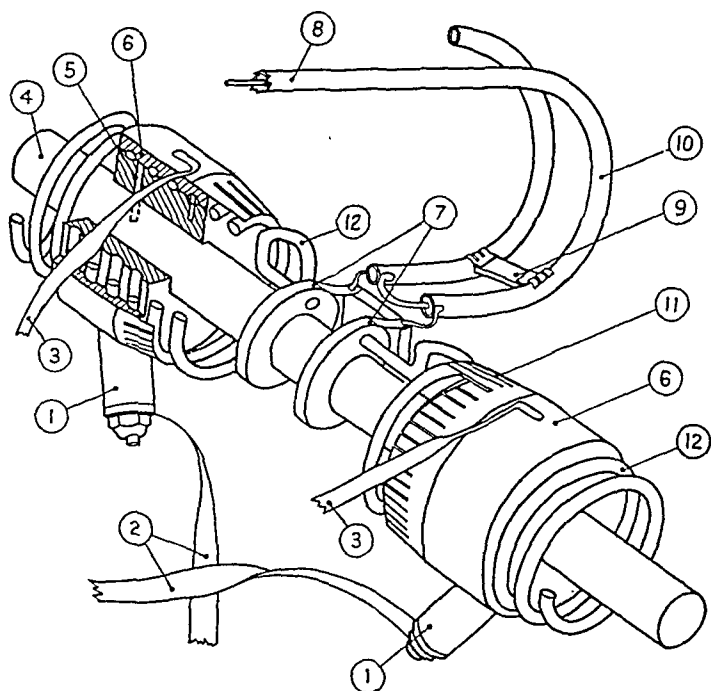


FIG. 14-25.—Variable-inductance tank circuit using short-circuiting sleeves: (1) feedback capacitors ($6\mu\text{f}$); (2) flexible leads to 829B grids for feedback; (3) flexible leads to 829B plates through d-c blocking capacitors; (4) shaft (rotates); (5) MF Formica coil form pinned to shaft; (6) movable sleeve, free to move axially when driven by rotation of coil but kept from rotation on tongue and guide; (7) slip rings and brush contacts; (8) 50-ohm line to load; (9) balun short-circuiting bar (adjustable); (10) conductively coupled balun output coupler; (11) finger contacts; (12) rotating coil (opposite winding direction in two halves); Note: All coils, slip rings, sleeves, and contacts silver-plated.

The parts, such as sides and covers, of the metal boxes used to shield oscillators should be well bonded together either by screws or by contact fingers. This is because electromagnetic shielding depends upon the flow of induced currents in the metal of the shield. These currents, by Lenz's law, produce magnetic fields that oppose the fields of the primary currents and cause the net field, outside the shield, to be weak. If these induced currents are weakened by poor contacts in the shield, the effectiveness of the latter will be reduced. For the same reason the shield should be constructed of a material of high conductivity. For ultrahigh

frequencies, silver is desirable. Since the depth of penetration of the current is small, silver plating is sufficient.

Shield boxes should in general fit the oscillator fairly closely. If any of the dimensions are $\frac{1}{2}$ wavelength long at any frequency within the range of the oscillator, there is danger of cavity resonances which will dissipate power and thus result in *holes*, or frequency ranges in which the power output is zero or very low. There is also the possibility of waveguide modes of resonance when a circumference of the shield box is equal to 1 wavelength. These modes are discussed in Sec. 15-26.

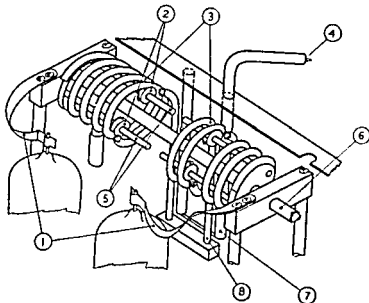


FIG. 14-20.—Variable-inductance "rollo" coil with short-circuiting contacts mounted inside the coil: (1) plate leads of amplifier tubes; (2) rollo contacts pinioned on spring hairpin and free to slide axially on hairpin; (3) fixed coils (opposite winding direction in two halves); (4) 50-ohm line to load; (5) rotatable hairpin spring assembly; (6) rotatable shaft; (7) inductively coupled balun output coupler; (8) center of tank inductance formed into loop for coupling to antenna-circuit balun.

Rods inside the oscillator shield box, such as feed screws, grounded at one end but not at the other, may be expected to give trouble from power-absorbing resonances if they are $\frac{1}{4}$ wavelength long at any frequency within the range of the oscillator.

14-12. Cathode Chokes and Tuned Lines.—The inductance of filament leads, which are of necessity relatively long, may be troublesome in grounded or common-cathode oscillators, as discussed in Sec. 14-1. It is not possible, at very high frequencies, to ground these filaments at the tube terminals. They may, however, be brought electrically to ground by means of a half-wave linear transformer, either parallel-line or coaxial, or by using relatively long small-diameter single-layer resonant chokes.

TABLE 14-1.—TYPICAL CHARACTERISTICS AND PERFORMANCE OF SEVERAL ULTRAHIGH-FREQUENCY LOW- AND MEDIUM-POWER TRANSMITTING TUBES

Tube	Type	g_m , μmhos	μ *	C_{sp} , μuf	C_{ek} , μuf	C_{pk} , μuf	Max. rated dissipation P_n , watts	Preferred circuit	Typical test conditions as oscillators				
									Fre- quency Mc	Plate volt- age E_b , volts	R-f out- put P_o , watts	Plate effic. η_p , %	Cooling method
3C24	Triode	2,500	23	1.5	1.7	0.3	25	Spiral-induct- ance tank circuit	100 320	750 750	50† 35†	40 22	Forced air
388A	"Doorknob" triode	2,000	6	1.5	1.5	1.0	50	Parallel line	200	550	20†	14	Forced air
368AS	"Doorknob" triode	2,500	8	1.1	0.9	0.6	20	Parallel line	550	550	5.5†	4	Forced air
2C29‡,§ (X102G)	Disk triode	5,500	...	3.2	5.2	0.12	30	Coaxial	450	425	8†	15	Forced air
3C22†	Planar triode "light- house"	5,000	45	2.4	4.9	0.08	125	Coaxial	720	425	3†	6	Forced air
									600	500	12	30	Forced air
									1250	500	7	12	Forced air
									500	950	45	40	Forced air
									1000	950	35	29	Forced air
									1500	950	15	12.5	Forced air
2C38†	Planar triode invert- ed "lighthouse"	20,000	100	1.9	6.5	0.04	100	Coaxial	300	925	55	55	Forced air
									1000	925	40	40	Forced air
									2500	925	15	15	Forced air
									200	1500	500	60	Forced air
									600	1500	200	40	Forced air
L600N	Triode	13,000	8	6.0	6.7	0.5	600	Coaxial	1000	1500	50	10	Forced air

* Average values, sometimes including early experimental tubes.

† Power output for two tubes in push-pull.

‡ Illustrated in Fig. 15-2.

§ Only experimental models tested.

|| Early type illustrated in Fig. 15-2

In the latter case the choke, acting as a loaded transmission line, is physically much shorter than the same electrical length of straight line.

In many oscillators, *e.g.*, those shown in Figs. 14-14, 14-16, 14-17, and 14-21, chokes are used to prevent r-f currents from flowing into the filament leads. Chokes used in this manner are common impedances in the grid and plate resonant circuits. Untuned chokes have the advantage of simplicity over tuned chokes in that no tuning controls are required. For this reason they are useful for single-dial wide-range portable equipments. However, since the filament end of a fixed choke cannot present either a high or a low impedance at all frequencies over a wide range, compromises must be made in performance, as compared with that of an oscillator using tunable cathode impedances. Over a wide tuning range, say 3-to-1, the impedance of the cathode chokes goes through a series of ranges such that the chokes may be equivalent to a quarter-wave line at some frequencies and to a half-wave line at others. The behavior of such an oscillator is therefore electrically very complicated. In practice, the cathode chokes are usually selected by trial and error so as to give as nearly as possible the desired performance over the required frequency range. By proper choice of filament chokes in a common-cathode type of oscillator, it is possible to obtain satisfactory operation over a wide range of frequencies without recourse to the mechanical complexity of tuned-cathode circuits, even though the chokes may satisfy the requirements for optimum performance over only a narrow range.

CHAPTER 15

COAXIAL-LINE POWER AMPLIFIERS AND OSCILLATORS

BY W. R. RAMBO

15-1. Coaxial-line Oscillators and Amplifiers.—The adaption of conventional oscillator and amplifier circuits to u-h-f use is facilitated by the use of coaxial lines as circuit elements. The high inherent Q of concentric lines as resonant circuits, the very low radiation, and the possibility of isolation of the circuits contribute to successful design.

A typical coaxial-line circuit element is shown in Fig. 15-1, and terms used in connection with it are defined in Table 15-1. One or more such lines may be used in an oscillator or amplifier to perform the circuit functions normally assigned to the parallel-wire lines or the lumped-coil and capacitor elements in the more familiar units.

TABLE 15-1.—DEFINITIONS OF TERMS AND FORMULAS USED IN CONNECTION WITH THE COAXIAL-LINE CIRCUIT ELEMENT OF FIG. 15-1

Z_0 , characteristic impedance of the line, ohms	$Z_0 = 60 \ln \frac{d_2}{d_1}, \dagger$
d_2 and d_1 are in any consistent units	$= 138 \log_{10} \frac{d_2}{d_1}$
L , inductance of the line, henrys per meter	$Z_0 = \sqrt{L/C}$
C , capacitance of the line, farads per meter	
v , velocity of propagation, 300×10^6 m per sec for an air-dielectric line	$v = \frac{1}{\sqrt{LC}}$
Z_{in} , line input impedance, ohms	$Z_{in} = jZ_0 \tan \theta$ (receiving end short-circuited)
θ , angular length of line, degrees	$\theta = \frac{360l}{\lambda}$
λ , wavelength along the line, meters	$\lambda = v/f$
f , frequency, cycles per second	$L = Z_0/v$
l , length of the line, meters	$C = 1/vZ_0$

* These formulas apply to uniform dissipationless lines. They are good approximations when the unloaded Q is very much higher than the loaded operating Q of the line used as a circuit element.

† If the line is filled with a dielectric of constant ϵ , multiply Z_0 as calculated above by $1/\sqrt{\epsilon}$.

15-2. Vacuum Tubes for Coaxial Circuits.—The difficulty of physically incorporating conventional u-h-f tubes (acorn tubes, doorknob tubes, etc.) in coaxial-circuit configurations has led to the development of a series of tubes specifically designed for such use. While the internal elements of these tubes may be either cylindrical or of the parallel-plane

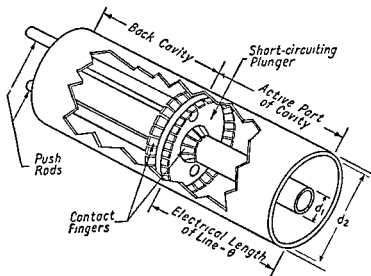


FIG. 15-1.—Coaxial-line circuit element

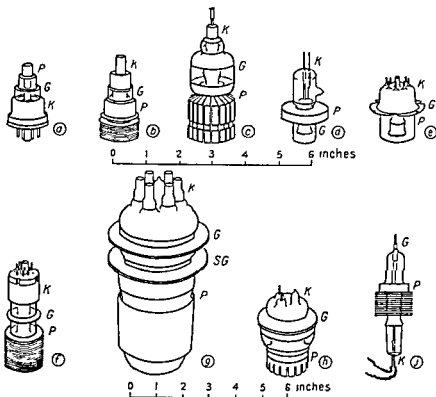


FIG. 15-2.—Typical tubes for coaxial-circuit use: (a) 2C44, (b) 2C38 (2C30), (c) A2214, (d) CV178, (e) X102G, (f) 3C22, (g) X139-D, (h) L600E, (i) N.E. Type 1.

type, all are characterized externally by the provision for a circumferential connection of circuit element to tube terminal. Thus, the conventional pin connectors are replaced by disks or flanges that are sealed to the tube envelope, to which a continuous, or multipoint, connection can be made from the cylindrical coaxial-line element through an appropriate connecting medium, generally a set of flexible fingers.

Typical tube types suitable for use in coaxial circuits are shown in Fig. 15-2.¹ These are triodes and tetrodes designed for powers ranging from a fraction of a watt to the kilowatt level, and for upper frequency limits from a few hundred to several thousand megacycles.

The cylindrical or disk construction is carried through from the external terminal of the tube to the actual active part of the tube elements. A high degree of circuit isolation is thus possible, and coupling between circuits is reduced to a minimum. Spacing of elements, lead inductance, and interelectrode capacitance are all subject to considerable control through these types of construction.

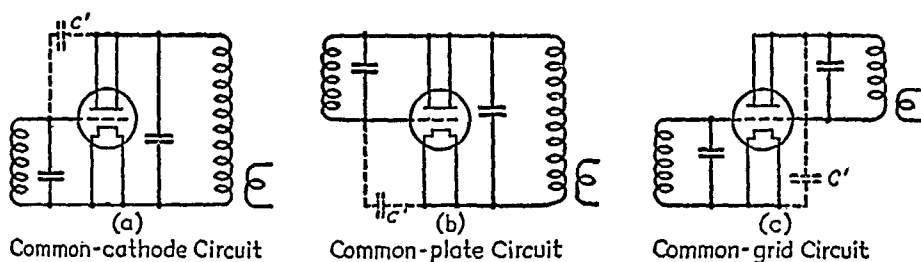


FIG. 15-3.—Oscillator and amplifier circuits suitable for coaxial-circuit use: (a) common-cathode circuit; (b) common-plate circuit; (c) common-grid circuit. C' is the tube inter-electrode feedback capacitance.

15-3. Circuit Choices for Ultrahigh-frequency Use.²—Many conventional oscillator and amplifier circuits, such as shown in Fig. 15-3, are suitable electrically for u-h-f use. Practically, the final choice of circuit is almost always a common-grid or grid-separation type of circuit in which the tube element common to the input and output circuits is the grid.

Common-plate circuits and the familiar common-cathode circuits are generally ruled out by the structure of this type of tube, the cathode-grid-plate order of electrode terminals pointing to the use of the grid as the common element. This is true both of triodes and of tetrodes operating in coaxial circuits where it is possible to operate both grid and screen-grid at the same r-f potential and as the common terminals.

The use of the grid-separation type of circuit is particularly advantageous in amplifier operation. The feedback or coupling capacitance

¹ Characteristics of several of these tubes are detailed in Table 14-1.

² A general discussion of the properties of oscillator circuits is given in Chap. 14.

between output and input circuits is the plate-cathode capacitance, which is reduced to a minimum in most tubes suitable for coaxial-circuit use. Thus regeneration through interelectrode feedback is materially reduced by this shielding action of the grid. Even without neutralization, most tubes of this type will not oscillate without external feedback in any but the upper limit of their normal operating range when the plate circuit is loaded.

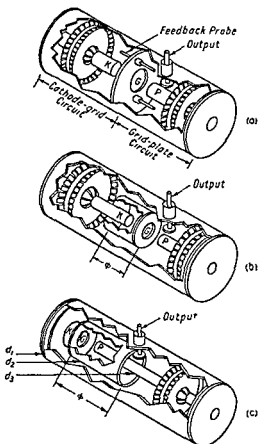


FIG. 15-4.—R-f coaxial circuits for common-grid operation (tube and tube contact fingers not shown). The feedback line has an electrical length ϕ and a characteristic impedance determined by the ratio of d_1 to d_2 .

The similarity between the grid-separation type of oscillator and amplifier circuits is considerable. The conversion of an oscillator to an amplifier consists primarily of removing the external feedback system, the addition of a source of driving energy, and retuning.

15-4. The Common-grid Circuit.—Figure 15-4 shows the basic common-grid circuit in several configurations suitable for oscillator operation. Two independently tuned circuits are provided in Fig. 15-4a, and relatively independent adjustments of circuit tuning and loading are

possible. Normal feedback of energy from output to input circuits through the tube plate-cathode interelectrode capacitance is usually augmented by an appropriate external feedback system such as the probe-coupling arrangement shown in Fig. 15-4a.

A modification of this coupling method is illustrated in Fig. 15-4b, in which the driving energy is transferred from the plate-grid circuit to the grid-cathode circuit through a coaxial feedback transmission line ϕ . Variation in driving-voltage phase and amplitude is possible through

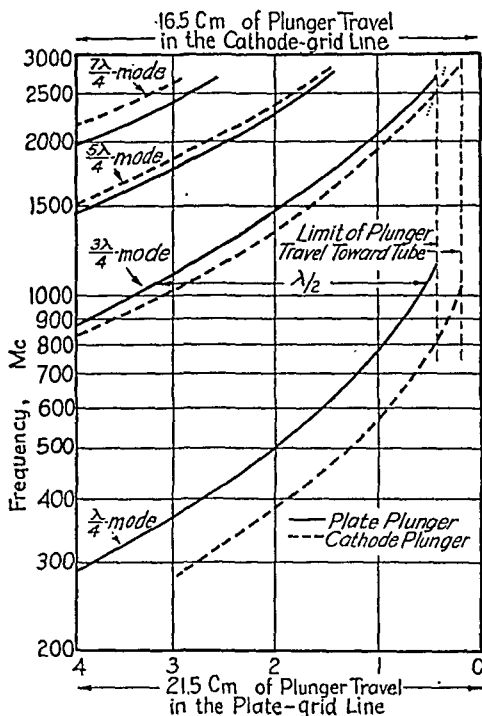


FIG. 15-5.—Tuning chart for a typical wide-range oscillator.

adjustment of the electrical length of the feedback line and of its characteristic impedance and that of the grid-cathode line.¹ This is an efficient way of transferring large quantities of feedback energy as needed under special operating conditions. It has the disadvantage of introducing an element, the line ϕ , that is electrically frequency sensitive and that may limit mechanically the motion of the tuning plungers. The introduction of grid bias necessitates the use of r-f chokes.

A third form is that shown in Fig. 15-4c, in which the independence of the two tuned circuits has been partly eliminated, frequency and excita-

¹ For single-frequency amplifier operation, it is possible to adjust amplitude and phase of the feedback energy to oppose that transferred through the tube interelectrode capacitance and thereby obtain a neutralizing action.

tion conditions being controlled by the adjustment of a single plunger and the characteristics of the line ϕ . Conditions in both circuits are affected by adjustments in either and in the output coupling circuit, and optimum operation is restricted to a narrow range around the design frequency.

When a wide tuning range is a design requirement, it is necessary to resort to the basic circuit of Fig. 15-4a. Independent adjustment of driving and output circuits allows, through proper tuning, the setting up of correct operating conditions over very wide frequency ranges. It permits, too, a further increase in tuning range through multiple-mode operation (tuning conditions in which the electrical length of one or both lines may be any odd multiple of $\frac{1}{4}$ wavelength). With this configuration, there is little limitation on tuning range through mechanical restrictions on tuning-plunger motion. The operation of the three circuits is comparable at a single design frequency. Something of a physical problem is posed in the circuits of Figs. 15-4b and c by the need for supporting the grid cylinder and for providing for ease of tube insertion and removal.

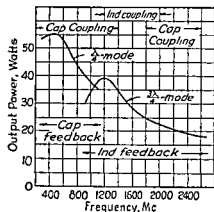


FIG. 15-6.—Power output as a function of frequency for a type-2C38 inverted lighthouse triode in a wide-range oscillator circuit.

largely a function of tube performance, possibly modified over incremental ranges by the characteristics of the feedback and output coupling systems. A typical output curve is shown in Fig. 15-6.

As in conventional circuits, the operating frequency is determined largely by plate-circuit tuning, while the power output and efficiency will depend on grid-circuit tuning, which affects amplitude and phase of the driving voltage. The plate-circuit loaded Q will influence both the frequency stability of the oscillator and the modulated band width and, for a given loaded resonant impedance, will depend on line dimensions, tube capacitance, and the operating mode.

Because of its wide application, the common-grid circuit of Fig. 15-4a will be discussed more fully in the following sections. As an oscillator, such a circuit is capable of tuning over as much as a 10-to-1 frequency range. This is illustrated in Fig. 15-5. The principal limitations on tuning range are imposed by the physical size of the circuit elements at low frequencies and by tube capabilities at the high-frequency end. The curve of output power vs. frequency is

15-5. Multiple-mode Tuning.—Coaxial-line tuned circuits are formed by sections of concentric line of such characteristic impedance Z_0 , and electric length θ , as to resonate or nearly resonate the capacitances terminating these lines. These are frequently tube interelectrode capacitances. The lines are generally short-circuited by movable plungers or pistons, the plunger motion providing an adjustment of electrical length. The characteristic impedance is fixed by the initial design.

By assuming a short-circuited lossless line of uniform cross section, the input impedance can be calculated from the expression

$$Z_{in} = jZ_0 \tan \theta \quad (15-1)$$

where Z_{in} = input impedance, ohms

Z_0 = characteristic impedance, ohms

θ = line length, degrees

The value of Z_{in} will not be changed through the addition, physically, of extra half-wavelengths of line

$$[\tan \theta = \tan (\theta + N \times 180 \text{ deg}), \\ \text{where } N = 1, 2, 3, \dots]$$

This is an important property for, in the practical sense, the electrical length of line within the tube envelope from the active portion of the tube element to the tube terminal is sufficiently great, at very short wavelengths, to limit seriously the minimum value of θ (and Z_{in}) achievable by plunger motion. By moving the physical short circuit $\frac{1}{2}$ wavelength away from the tube, it is possible to move the effective electrical short circuit actually into the tube itself and thereby resonate the small reactance of the interelectrode capacitance with the very short length of line to the first electrical short circuit.

This leads to the common practice in circuits of this type of so-called

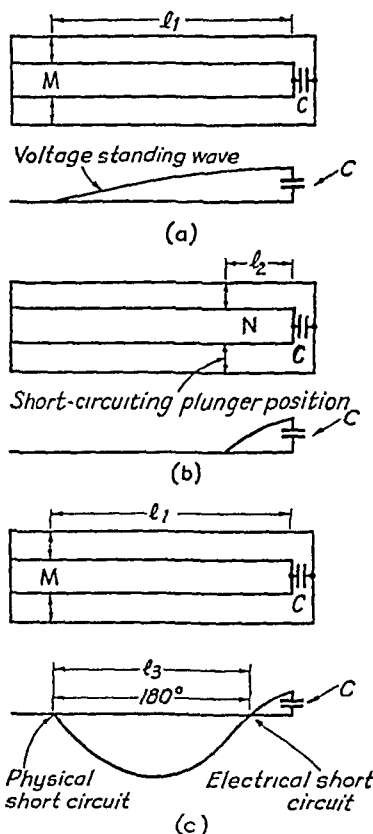


FIG. 15-7.—Tuning conditions in a resonant coaxial line. (a) $\lambda/4$ mode operation at low frequency f_1 . (ℓ_1 is less than 90° at f_1). (b) $\lambda/4$ mode operation at medium frequency f_2 . (c) $3\lambda/4$ mode operation at medium frequency f_2 (ℓ_1 is less than 270° but greater than 180° at f_2).

multiple-mode operation for obtaining wide tuning ranges.¹ The low-frequency limit of a line in $\lambda/4$ mode operation is that at which, with the maximum physical length included in the circuit (as l_1 in Fig. 15-7a), the input impedance as given by Eq. (15-1) (with $\theta < 90$ deg) approximately equals the terminating capacitive reactance. The resonant frequency increases as the short-circuiting plunger is moved toward the tube, the upper frequency of the $\lambda/4$ mode of operation probably being limited by the mechanical motion of the plunger. The plunger is then moved an electrical half-wavelength away from the tube at the frequency in question. The circuit again resonates, but now with an electrical length θ' such that $180 \text{ deg} < \theta' < 270 \text{ deg}$. The electrical length of the line in degrees falls in the third quadrant, and so $3\lambda/4$ mode operation results. A motion of the plunger toward the tube will increase the operating frequency above that obtainable on the $\lambda/4$ mode and, by this maneuver, two or more frequency ranges may be tuned by the plunger in traversing the same physical section of line. Each higher range is covered in a higher order tuning mode, as shown in Fig. 15-7, *i.e.*, the electrical length of the line at resonance equals $\lambda/4$, $3\lambda/4$, $5\lambda/4$, etc.

15-6. Mode Separation.—The ability of a coaxial line to resonate with a terminating capacitance simultaneously at several frequencies not harmonically related presents a choice of operating frequencies to an oscillator. That which it chooses will be the one for which proper tuning conditions exist in both input and output circuits regardless of the operating mode.

It is possible that tuning conditions will be satisfied for more than one frequency simultaneously in both circuits with a given plunger setting. This presents a choice of frequencies to an oscillator and erratic operation may manifest itself in "dead" spots, regions in the frequency range in which no oscillations are present, or in a tendency of the frequency to jump from one range to another. The final choice is that permitting highest Q operation, which is a function of frequency and relative loading. The remedy is to design the two tuned circuits so that, with a given plunger setting, the tuning conditions for oscillation are satisfied at only one frequency.

Proper circuit design for good "mode separation" lies in making sure that the relative plunger settings producing tuning conditions at a frequency in one mode of operation do not also tune the circuits to another frequency in another mode.

The conditions leading to the worst possible case of lack of mode separation is that in which both input and output circuits consist of equal

¹ "Mode" as used in this sense refers to the number of electrical quarter wavelengths in the resonant-line voltage standing wave rather than to the "resonance mode" determined by the electromagnetic field pattern.

terminating capacitances tuned by lines of equal characteristic impedance. Then

$$\frac{C_1 Z_{01}}{C_2 Z_{02}} = 1 \quad (15-2)$$

where Z_{01} = characteristic impedance of the input line, ohms
 Z_{02} = characteristic impedance of the output line, ohms
 C_1 = terminating capacitance of input line
 C_2 = terminating capacitance of output line

Both circuits would then always be simultaneously tuned at several frequencies in several modes when plunger positions are such as to produce oscillations of the desired frequency. Since the terminating capacitances are usually the interelectrode capacitances of vacuum tubes and are generally dissimilar, this worst case rarely occurs. More often, the result of poor mode separation is a single narrow band of frequencies in a given tuning range in which it is impossible to achieve stable operation. A clue to the remedy in such cases is given in Eq. (15-2), which represents the worst circuit conditions. By designing the circuit so that the product $C_1 Z_{01}$ is as different as possible from $C_2 Z_{02}$, good mode separation should be assured.

Table 15-2 was prepared from the wide-range oscillator tuning characteristics plotted in Fig. 15-5 and shows several possible tuning combinations in the grid-cathode and grid-plate circuits. The plate plunger was moved away from the tube until a considerable length of line lay between the tube and plunger to allow operation on high-order modes at high frequencies. The plate plunger was left fixed in this position (at a relative setting of 3.5 in Fig. 15-5). The cathode-circuit plunger was advanced toward the tube from the extreme end of the grid-cathode cavity (relative position 4 in Fig. 15-5), and the frequency and tuning

TABLE 15-2.—TYPICAL TUNING CONDITIONS IN A WIDE-RANGE OSCILLATOR*

Setting	Cathode-plunger position	Oscillator frequency produced, Mc	Tuning mode in plate-grid line	Tuning mode in cathode-grid line
A	4.00	2150	$7\lambda/4$	$7\lambda/4$
B	3.75	1600	$5\lambda/4$	$5\lambda/4$
C	3.10	1000	$3\lambda/4$	$3\lambda/4$
D	2.50	330	$1\lambda/4$	$1\lambda/4$
E	2.35	2150	$7\lambda/4$	$5\lambda/4$
F	1.50	1600	$5\lambda/4$	$3\lambda/4$
G	0.70	2150	$7\lambda/4$	$3\lambda/4$
H	0.20	1000	$3\lambda/4$	$1\lambda/4$

* With plate-plunger position fixed at a relative setting of 3.5 (see Fig. 15-5).

modes were checked at each cathode-circuit plunger position producing oscillations. These data are recorded in Table 15-2.

The first four critical settings of the cathode plunger produced oscillations at widely separated frequencies as both lines resonated in successively lower modes. The total cathode-plunger motion between settings *A* and *D* was only 6.2 cm. The tuning-mode conditions in the plate line remained the same in settings *A*, *E*, and *G* while the cathode line resonated at 2150 Mc in the $7\lambda/4$, $5\lambda/4$, and $3\lambda/4$ modes, respectively, the plunger traversing $\frac{1}{2}$ wavelength of line between settings.

There are obviously many plunger-setting combinations that will produce oscillations in such an oscillator, and dials or scales must be incorporated in a unit to permit resetting of tuning conditions and to assist in the determination of operating mode through relative plunger settings.

Poor mode separation would appear in Table 15-2 as a lack of separation of values in the column of cathode plunger settings. Interchangeable high- and low-frequency feedback units were used in this oscillator which strongly favor one or another mode and range of frequencies and aid materially in the suppression of mode interference. Thus, oscillations were detected only at settings *A*, *B*, *C*, *E*, *F*, and *G* in Table 15-2 with the high-frequency unit in place, while the low-frequency feedback suppressed all oscillations but those occurring at settings *D* and *H*.

15-7. Loaded-*Q* Considerations.—Whereas in the ideal case the expression for the input impedance of the coaxial line is frequently treated as a pure reactance, it should not be forgotten that the line actually is a circuit element with distributed constants, both inductive and capacitive. While the inductive reactance of a short-circuited section of line less than 90 deg in length may be used to tune out a terminating capacitive reactance, the total capacitance in the resonant circuit is materially increased by that distributed in the line.

The distributed capacitance of a coaxial line is a function of the characteristic impedance. This is of importance in the choice of line characteristic impedance for a tank circuit where high operating *Q* must be considered for its limitation on the modulated band width or, in the case of an oscillator, for its influence on frequency stability.¹ A given input reactance might be obtained with a short high-characteristic-

¹ The discussion of desirable plate circuit operating *Q* in this chapter is based on the assumption that wide-band modulation is a design requirement necessitating the maintenance of as low a loaded *Q* as possible for a given shunt impedance. If oscillator frequency stability were of more importance, a high-operating *Q* would be desirable and would require an attendant modification in circuit design. The properties of coaxial lines as frequency-stabilizing circuit elements have been discussed elsewhere. For example, see F. E. Terman, "Radio Engineers' Handbook," p 435, McGraw-Hill Book Company, Inc., New York, 1943.

impedance line or a long low-characteristic-impedance line. The resonant-circuit Q of the short line, when shunt loaded with a given resistance, will be lower than that of the longer line if the electrical length of the lines is less than 90 deg. Either of the lines of Fig. 15-8 may be used to resonate a given terminating capacitance, but for a given impressed voltage and resonant impedance, as determined by loading, the extra storage of energy in the low-impedance line (Fig. 15-8a) will increase its operating Q over that of the high-impedance line (Fig. 15-8b).

This discussion assumes that the internal losses of the lines as circuit elements may be neglected, *i.e.*, that the unloaded Q is very much higher than the loaded Q when the circuit is coupled to an external load. It assumes, too, that the loaded impedance at resonance is fixed, being that value which properly matches the impedance of the tube. In actual practice, it is found that the loaded tank-circuit Q of typical coaxial-line oscillators may run as high as 300 to 800 in higher mode operation so that when wide-band modulation is a requirement it is necessary to consider all means by which it can be reduced. It is then normal practice to employ high-characteristic-impedance lines for circuits operating in the $\lambda/4$ mode.¹

Higher mode operation introduces complications in that loaded Q , considered on the basis of energy stored and energy delivered per cycle, is materially increased by the additional storage of energy in the added half-wavelength sections of line. Thus, at a given frequency, the loaded Q of the line in Fig. 15-9b is much greater than that in Fig. 15-9a, even though the impedance at resonance may be the same. While the switch to higher mode operation will invariably increase loaded Q , the question again arises as to what is the best line-characteristic impedance to hold this increase to a minimum. It was concluded that a high-characteristic-

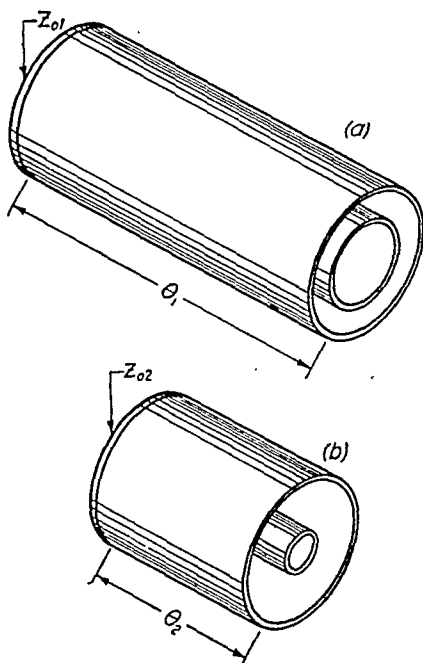


FIG. 15-8.—Coaxial tuning elements: (a) low characteristic impedance; (b) high characteristic impedance.

¹ The effect of circuit Q on tubes modulated over wide band widths is discussed in Chap. 17. It is generally possible to operate amplifier plate circuits more heavily loaded and with a lower Q than oscillator plate circuits so that wider modulated band widths are possible.

impedance line was desirable for $\lambda/4$ mode operation, but generally this is not true for higher mode operation. This difference can be explained by recalling that the actual stored energy is a function not only of the line capacitance, but also of the square of the voltage across the capacitance. The voltage standing-wave distribution in a resonant line is a section of a sine wave, and the total stored energy must be calculated by a process of summation of the contributions of incremental sections along the line.

In higher mode operation, a change of characteristic impedance affects the total stored energy and Q , not only through the line capaci-

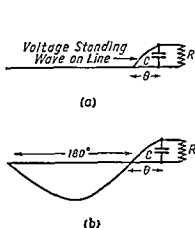


FIG. 15-9—Voltage standing waves on short-circuited resonant coaxial lines operating in (a) the $\lambda/4$ mode and (b) the $3\lambda/4$ mode. The added stored energy in the additional 180 deg of line in the $3\lambda/4$ mode increases the resonant-circuit Q when the line is loaded by a fixed resistance R . The line-terminating capacitance is C . $\theta < 90$ deg.

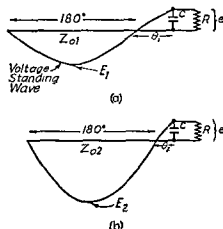


FIG. 15-10—Effect of coaxial-line characteristic impedance on resonant-circuit voltage. The lines resonate with a fixed capacitance C , and are loaded by R to a fixed resonant impedance. $Z_{01} < Z_{02}$; $\theta_1 > \theta_2$; $E_1 < E_2$; $Q_1 < Q_2$, where E_1 and E_2 are the maximum line voltages, θ_1 and θ_2 , the electrical lengths to the first electrical short circuit.

tance, but also, and in an opposite way as far as the Q is concerned, by a variation in the average line voltage. In Fig. 15-10 it is assumed that a resonant line is driven by an r-f voltage e as determined, for example, by the plate swing of an oscillator. Two cases of $3\lambda/4$ mode operation are considered, that of a low characteristic impedance and that of a high characteristic impedance. The resonant impedance is the same in either case, the assumed lossless line being loaded by the parallel resistance R .

The low-characteristic-impedance line (Z_{01}) will require that θ_1 be larger than θ_2 if both lines are to resonate with C_{sp} . Since both lines are driven by the same voltage e , it is apparent that E_2 is larger than E_1 since

$$e = E_1 \sin \theta_1 = E_2 \sin \theta_2$$

or

$$E_2 = E_1 \frac{\sin \theta_1}{\sin \theta_2} \quad (15-3)$$

and

$$\sin \theta_1 > \sin \theta_2$$

This increase in voltage in the added half-wave sections of line by the decrease in characteristic impedance frequently increases the stored energy more than the reduction in capacitance per unit length will reduce it. In the practical case, it is almost always possible to reduce the loaded Q by a reduction in characteristic impedance when operating in higher modes, for the loaded Q will be reduced whenever θ_2 , as in Fig. 15-10*b*, is less than about 40 deg with $3\lambda/4$ mode operation, and about 25 deg with $5\lambda/4$ mode operation. With the normal run of interelectrode capacitances of the small power tubes, and with the characteristic impedances of coaxial lines that are physically practical to construct, θ_2 may run anywhere from 45 to 10 deg or less at frequencies above 1000 Mc.

It can be concluded, then, that if a circuit is to be operated in the $\lambda/4$ mode, a high characteristic impedance is advantageous in reducing operating Q . If all or most of the operating range must be covered in higher modes, a low-characteristic-impedance line is more practical. Insofar as physical dimensions are concerned, "low" and "high" would probably be limited to about 20 and 90 ohms, respectively. If the range must be covered in both $\lambda/4$ and higher modes, it will be necessary to effect a compromise in characteristic impedance, and, indeed, such a compromise may be dictated by other considerations governing the choice of dimensions.

15-8. Limitations on Tuning Range.—A practical limitation on the low-frequency range of a coaxial oscillator or amplifier is the actual physical length of the line elements, which rapidly increases as the frequency is lowered. This can be appreciated when the actual physical quarter wavelength is considered at low frequencies, for the resonant lines approach this length quite closely as the reactance of a fixed terminating capacitance increases with the decrease in frequency. When over-all physical length is an important consideration, it is helpful to remember that a given terminating capacitance may be tuned, with a fixed maximum length of line, to a lower frequency with a line of higher characteristic impedance.

Physical dimensions also influence the practicable upper frequency limit of coaxial lines as resonant circuit elements, because of the ability of cavities of large radial electrical dimensions to support interfering waveguide and spurious coaxial resonance modes, "mode" in this sense referring to the alternating electromagnetic field pattern rather than the

number of quarter wavelengths in the voltage standing wave along the line. Generally, it is possible to suppress such mode interferences by avoiding electrical and physical designs that favor their existence. Tube operating characteristics then become the principal limitation to the high-frequency range. The principal interfering higher order coaxial resonance mode is the $TE_{1,1}$ mode, which can exist only at wavelengths less than the cutoff value given by

$$\lambda_c \approx \pi(r_1 + r_2) \quad (15-4)$$

where λ_c is the cutoff wavelength, r_1 the radius of the inner conductor, and r_2 the radius of the outer conductor.

In any event, this $TE_{1,1}$ mode cannot interfere if the resonant-circuit line lengths are less than 90 deg.¹ A plot of resonance wavelength vs. physical line length for the $\lambda/4$ and $3\lambda/4$ TEM modes (normal operating

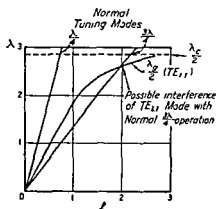


FIG. 15-11.—Coaxial-line resonant-circuit lengths for several possible tuning modes.

impedance and requires, too, an efficient system for the transfer of energy from the output to the grid circuit.

This transfer of energy, in the simplest case, may take place entirely in the interelectrode tube capacitance, but this capacitive feedback is generally augmented by a feedback system external to the tube. Perhaps the simplest and least frequency-sensitive scheme is merely to link output and input circuits through a common wall, inductively by small loops or capacitively by insulated probes. Wider frequency ranges may be obtained by a combination of the two. Other common methods are

¹ While the $TE_{1,1}$ mode may exist in the line at any wavelength less than the cutoff wavelength given by Eq. (15-4), interference usually occurs only if the mode is propagated at a frequency at which the coaxial-line length is $\lambda_c/2$ and is simultaneously at the right electrical length so that the line resonates in a normal TEM mode. As shown in Fig. 15-11, this can happen only with the higher order TEM modes.

modes) and the $TE_{1,1}$ interfering mode appears in Fig. 15-11. As the operating wavelength approaches the cutoff wavelength, the $TE_{1,1}$ guide wavelength approaches infinity and there will be an intersecting point with the $3\lambda/4$ and higher normal operating modes at which erratic operation can exist.

15-9. Feedback Considerations.

Optimum oscillator operation requires that a grid driving voltage be supplied in both proper amplitude and phase. This is influenced somewhat by grid-circuit tuning and the grid-circuit

the internal transmission-line arrangements of Figs. 15-4*b* and *c* and a coupling of the two circuits through adjustable external transmission lines to provide an external phase adjustment.

In common-grid circuits, the wall common to both input and output circuits is that separating the two in the plane of the grid, as shown in Fig. 15-12*a*, or the cylinder forming the outer conductor of one concentric line and the inner conductor of the other, as shown in Fig. 15-12*b*. Feedback energy is transferred through this common wall by feedback loops and probes.

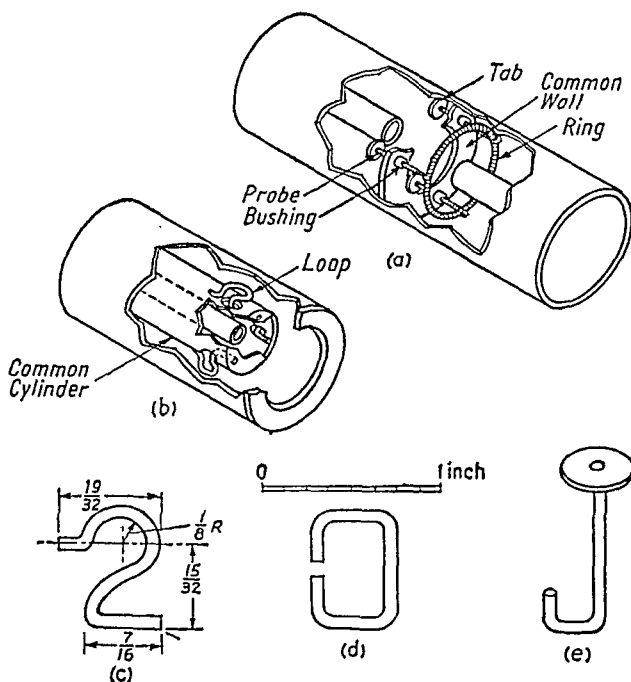


FIG. 15-12.—Coaxial-line oscillator untuned feedback systems. Tubes and tube contacts not shown.

It is generally advantageous to place either the loops or the probes, or both, physically and electrically as near the plane of the grid as possible, for in this region, the change in voltage or current amplitudes (depending on the frequency range) associated with changes in the standing-wave maximums and minimums with changing frequency is small, as shown in Fig. 15-13. The more nearly constant fields tend to make the amplitude of feedback more nearly independent of frequency. An additional advantage is that the feedback elements are removed from the region of plunger travel.

A few general statements can be made about the choice of feedback. When $\lambda/4$ operation is contemplated, the tuning line used is sufficiently

long electrically at the low frequencies to ensure a considerable electric field near the tube in both circuits. A capacitive or probe-type feedback is therefore ordinarily used. This, in the commonest form, consists of wire probes passing through the separating wall in insulated bushings and terminating in the two cavities in small disks or tabs or in wire rings linking the ends of the probes. The amount of feedback is determined by the intensity of the electric field (determined by where the feedback unit is placed with respect to the standing wave voltages along the line), the size of tabs or ring, and the proximity of tabs or ring to the respective inner conductors of the coaxial lines.

It is generally true that inductive-loop feedback placed near the tube will not work satisfactorily for low frequencies, because of low current in this region at these frequencies. However, as the frequency is increased,

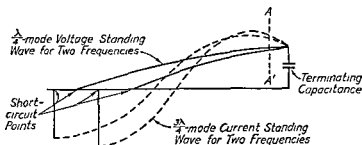


FIG. 15-13—Standing-wave plots in resonant coaxial line. As the position of the feedback elements ($A-A'$) approaches the terminating capacitance, there generally is less change in intensity in the $\lambda/4$ mode electric field and $3\lambda/4$ mode magnetic field.

the current maximum at the position of the electrical short circuit in the line approaches the plane of the feedback elements and loop coupling becomes very effective. With a still further increase in the frequency, for the $3\lambda/4$ or higher modes, the maximum-current point may pass through the feedback and approach the grid, but since it is usually possible to place the feedback loop physically close to the grid, loop coupling will remain effective to very high frequencies. This is because there is sufficient current density in the line for at least 45 electrical degrees on either side of the maximum point to operate properly the loop feedback. Such inductive feedback placed close to the grid plane, as in Fig. 15-12b, is effective over very wide frequency ranges and is particularly suitable to higher mode operation.

The most common forms of loop feedback are S loops, usually three in number, arranged symmetrically in the cavity. The open ends of the S are soldered to the inner and outer surfaces of the common conductor, the wire passing through a small hole or slot linking the cavities. Several small loops in a high-density magnetic field are more effective than fewer large loops in the same or less intense fields. Several modifications of

loop feedback are shown in Figs. 15-12*c*, *d*, and *e*. The phase of the feedback is reversed in the O loop from that in the S loop and the loop is generally useful over only a small frequency range. The J loop is designed to operate in a strong current field in one line and a strong voltage field in the other.

When an oscillator must cover a wide tuning range in both $\lambda/4$ and higher modes, there are two approaches to the untuned-feedback problem. The first consists of supplying both inductive and capacitive feedback elements so that, as the fields change in the vicinity of the feedback, one or the other will be effective. While this scheme is practical, it is complicated by the fact that, as the current standing-wave maximum passes through the feedback region, there is a reversal in phase in the electric field which, if it does not occur simultaneously in both cavities, will result in at least a narrow frequency range in which the transfer of energy through the probes is in the wrong phase to assist oscillation, the contribution of one type of feedback opposing the others.

A second method of obtaining wide-range feedback is to change the type of feedback at some convenient dividing frequency, *i.e.*, to remove the low-frequency feedback, perhaps probes, and to replace it with a high-frequency loop feedback.¹ A convenient method of making the change is to mount the feedback on a plate or short section of line, perhaps containing the grid-contact mechanism, that can be arranged to mount in the appropriate place in the circuit. Such an arrangement is shown in Fig. 15-18. The trouble of changing feedback is offset by several considerations.

1. Very wide useful ranges with good overlap can be arranged with each feedback so that the change need be made only infrequently.
2. More nearly optimum feedback can be provided over the entire range.
3. There is an important automatic suppression of undesired frequencies through the tendency of the feedback to favor one or another range. This action is a material aid in mode separation.

Most tubes will oscillate, even if only feebly, over much of their operating range with almost any form of feedback, as a considerable adjustment of driving-voltage phase and amplitude is possible through circuit tuning. It is, however, the function of a good feedback system to provide the best possible driving-voltage conditions over a wide frequency range.

15-10. Special Properties of Coaxial-line Tank Circuits.—In the interests of simplicity, the coaxial-line circuit elements discussed have

¹ While probes and loops were used for low- and high-frequency feedback, respectively, in the discussion, it is entirely possible that probe feedback could be used at high frequencies and loops at low frequencies if placed properly in the cavities. However, the arrangement described is generally the most effective.

been assumed to be of uniform dimensions throughout their effective electrical length. Discontinuities that do occur, principally in the portion of the circuit within the tube envelope or at the connection of the circuit element to the tube, generally do not destroy the qualitative validity of this assumption.

A smooth continuous line that tunes a terminating capacitance will not resonate in higher modes on harmonically related frequencies. A line formed of two series sections of different characteristic impedances and electrical lengths, however, can be made to form a circuit resonant simultaneously on a fundamental and on a harmonic frequency and leads to some interesting circuit possibilities. For example, oscillators have been constructed in which the tank circuit presents an impedance to both fundamental and third-harmonic currents. The circuit is loaded through a high-pass filter only on the harmonic frequency to provide a means of obtaining third-harmonic energy. A single triode tube then serves as a form of combination oscillator and tripler, although the action is inefficient and the harmonic power output is low.

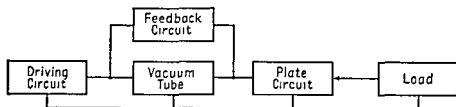


FIG. 15-14.—Block diagram of general oscillator circuit.

15-11. Ultrahigh-frequency Coaxial Oscillator and Amplifier Operation.—The differences between low-frequency and high-frequency coaxial oscillator and amplifier design and performance are not basic (the block diagram of Fig. 15-14 applies equally well to high- and low-frequency oscillators) but are the result, rather, of practical difficulties in the realization of desired circuit and tube properties. The plate-tank-circuit tuning and loading are arranged to give the impedance into which the power tube should operate under Class B or Class C conditions. The use of a high-current tube designed to operate into a low plate impedance tends to result in a reduction in operating Q , although this may be offset by an increase in stored energy in the circuit if the high-current properties were obtained at the expense of increased tube-output capacitance. Because of the high internal efficiency of coaxial lines as resonant circuit elements, a good ratio of unloaded to loaded Q , and hence high circuit efficiency, is possible even with the high loaded Q resulting from the large storage of reactive energy in coaxial circuits.

The grid circuit is tuned in a manner to set up proper driving-voltage phase and amplitude conditions and normally operates at a reasonably

low Q because of the high grid-cathode input conductance in shunt with the circuit. This conductance is large in a common-grid circuit because of the flow of r-f plate-cathode return current in the input circuit, which causes a transfer of energy directly to the plate circuit. An approximate analysis on the basis of 180-deg plate-current pulses shows this conductance to be as high as one-half the effective transconductance of the tube. The transconductance of typical u-h-f power tubes may easily be as high as 20,000 μ mhos and may be very much higher in pulsed operation. This low grid-circuit Q has two practical consequences. (1) The low driving-circuit Q permits considerable detuning in the circuit without serious effect on efficiency, so that plungers may be mechanically ganged to provide single-dial tuning over considerable ranges without grid-circuit trimming. (2) A large transfer of energy from output to input circuits is required to develop a driving voltage across the low grid-circuit impedance. Since most of this energy reappears in the output circuit as useful output, the over-all circuit efficiency is not impaired, provided that the feedback system is capable of handling the energy without undue loss in the coupling loops and probes.

As the operating frequency is increased, circuit losses increase, and the circuit efficiency drops. Increasing circuit losses and the poor tube efficiency contribute to a decreasing output and over-all efficiency. This is illustrated in Table 15-3 for the typical case of a small air-cooled power triode of the inverted lighthouse type (type 2C38). The low-frequency conditions were calculated from tube characteristics and are compared with measured performance at 1200 Mc in a common-grid oscillator circuit in $3\lambda/4$ mode operation.

TABLE 15-3.—COMPARISON OF CALCULATED LOW-FREQUENCY 2C38 OPERATION WITH MEASURED RESULTS AT 1200 Mc

	Frequency, Mc	E_{dc} , volts	I_{dc} , ma	P_{in} , watts	P_{out} , watts	Efficiency, %	Grid bias, volts	Peak grid volts	Grid current, ma	Angle of flow, deg	Plate swing, volts	Load impedance, ohms	Plate circuit Q
Calc.....	Low	1000	90	90	66	73	-25	+28	20	150	800	4000	...
Meas.....	1200	1000	90	90	40	44	-30	25	...	650	5200	250

At 400 Mc, the tube develops 54 watts at 60 per cent efficiency, and at 2500 Mc the power and efficiency have dropped to 15 watts and 17 per cent, respectively.

Figure 15-6 shows the general behavior of this tube as a plate-modulated Class C power oscillator in a circuit capable of being tuned over a

wide range. This performance was measured with $\lambda/4$ and $3\lambda/4$ mode operation only but, in the higher frequency range, it was possible also to check operation on the $5\lambda/4$ and $7\lambda/4$ modes. Over-all efficiency was only slightly impaired by the reduction in unloaded-to-loaded- Q ratios in the higher modes. The difference in output in the overlapping frequency range is due to feedback characteristics.

15-12. Grid-circuit Considerations.—The high operating frequencies of coaxial-circuit oscillators and amplifiers accentuate certain problems in connection with grid-circuit design. Grid bias may be obtained in any of the conventional ways; grid-leak bias, cathode bias, or a combination of the two are suitable. Cathode bias is an aid in preventing abnormally high plate currents in the high transconductance tubes when in the nonoscillating state but must be adjustable so that it can be lowered as a starting aid at very high frequencies. Grid-leak bias does not limit the nonoscillating plate current, but affords protection from destructive grid currents when arcs occur in u-h-f tubes with their close element spacing.

Intermittent operation or grid-circuit quenching is a fairly frequent occurrence at the high-frequency end of the tube range where the period of the r-f cycle is short compared with the normal grid-circuit time constant. Quench frequencies may run as high as several megacycles. The tendency to quench is reduced by high-plate-current and low-plate-voltage operation, by reducing grid-leak resistance and shunt capacitance, and in some cases by applying a fixed positive bias to the grid and increasing the amount of grid-leak resistance to achieve normal operating-bias conditions.

Grid emission is a source of trouble with some tubes. While this is a property of the tube, conditions can be improved by operating with low grid currents obtained through proper bias and loading adjustments.

Physical distortion or warping of the grid through excessive heating also has been noted. Under some conditions of operation, grid dissipation (plus heat radiated by the cathode) may be sufficiently high to produce a deformation of the grid structure in some tubes. This generally reacts on operating conditions to reduce the heating and permit the grid to cool and resume its normal position. The consequent repeating low-frequency cycle based on thermal conditions is manifested in unstable operation. These effects are less noticeable at the lower frequencies where changes accompanying the distortion will cause less reaction in over-all circuit conditions. Here, again, it is important to reduce grid dissipation to a minimum.

Cathode back heating, probably caused by dissipation of energy in electrons returning to the cathode through unfavorable transit-time conditions, is sometimes noticed in power tubes operating in the upper

section of their frequency range. For example, under some conditions, a type 2C39 tube operating at a plate input of 100 watts at 2500 Mc and with an efficiency of around 15 per cent may develop sufficient cathode back-heating to permit the complete removal of the normal filament power. A reduction of external filament power is then necessary to prevent abnormally high cathode temperatures and to restore emission to its normal value.

15-13. Circuit Tuning.—The principal tuning conditions peculiar to coaxial-line oscillators arise through the use of multiple-mode operation. The frequency of oscillation is determined largely by the setting of the plunger in the output or plate tank circuit. For a given tank-circuit tuning, there are usually several grid-circuit plunger settings that will produce oscillations. These oscillating frequencies may be identical, corresponding to operation in several modes in the grid line, or may differ from each other as grid-circuit tuning conditions pass through those permitting higher order resonances in the fixed length of tank-circuit line. Figure 15-5 illustrates possible frequencies of oscillation as functions of plunger settings in a typical coaxial oscillator. Table 15-2 lists possible operating frequencies and tuning modes in this oscillator for a certain fixed plate-circuit plunger position.

It will be found that tuning of the grid circuit is less sharp than that of the plate circuit, particularly in the lower frequency ranges. The tuning that produces the maximum amplitude of driving voltage is not necessarily that producing the best conditions for over-all operation. Frequently, a slight increase in grid-circuit line length will produce a driving-voltage phase adjustment that will more than offset loss in driving-voltage amplitude and will result in both more efficient and more stable oscillations.

It is good practice to select in both lines the lowest operating mode in which the desired frequency may be obtained. This is favorable from the standpoint of maintaining both a low loaded Q and a high inherent unloaded-circuit Q .

Care must be exercised in tuning coaxial oscillators, because the change in operating conditions resulting from the possible wide shift in frequency obtainable by a relatively small motion of the tuning plungers may produce damaging tube currents. This effect arises from the fact that, through multiple-mode operation, the plate circuit, for example, might resonate on widely separate frequencies through a small readjustment of grid-circuit tuning. A violent oscillating condition may result not only from the tendency to oscillate more powerfully at a lower frequency, but also because the new frequency might not be subject to the previous loading conditions. Whenever possible, preliminary tuning should be conducted at reduced plate voltages.

Tube reaction in the change from a nonoscillating to an oscillating state will depend on the frequency involved in relation to the capabilities of the tube. Generally, in the low-frequency range of the oscillator, the transition to an oscillatory state will result in a sudden rise in grid current and a decrease in plate current as would be expected. The application of load decreases the grid current and increases the plate current, and a normal loading procedure can be followed. At the high-frequency end of the range, and depending somewhat on the type of bias, a rise in both grid and plate currents often accompanies the start of oscillation. Loading at high frequencies may be applied with little reaction on plate current, and the maximum load is limited by the decrease of grid current and grid drive to the point of instability. It is generally found that maximum output is reached with a degree of loading just approaching this unstable point. Grid-current metering provides a valuable indication of operating conditions.

15-14. Amplifier Operation.—The common-grid coaxial circuit with all external feedback removed is particularly adapted to amplifier operation through the inherent tendency to suppress regeneration. The circuit requires large driving powers, but the interelectrode feedback and resulting tendency to oscillate is minimized by the small plate-cathode capacitance of tubes designed for coaxial use.

Without neutralization, the maximum usable frequency of the circuit for amplification is limited by the maximum allowable degree of regeneration. This frequency limit generally can be extended by operating the plate circuit heavily loaded. The resulting low plate-circuit Q also materially aids wide-band modulation.

For a given tube and frequency, oscillator and amplifier plate-circuit efficiencies and outputs are similar when allowances are made for the oscillator output going into the drive. The apparent power gain varies with the class of operation. For example, a lighthouse type of power triode, the 3C22, operating at 600 Mc in a circuit such as described in Sec. 15-27 develops a gain of 11 db in Class AB operation, 8 db in Class B operation, and 5 db in Class C operation. As a Class C amplifier, 75 watts output may be obtained with a plate input of 150 watts. Both gain and efficiency decrease as the operating frequency is increased.

A measurement of true gain in an unneutralized common-grid amplifier operating at ultrahigh frequencies is difficult to obtain, for a portion of the driving power is transferred directly to the output circuit. The apparent gain is a function of the regeneration present in the circuit and will depend on the unneutralized feedback capacitance, frequency, and operating conditions.

15-15. Practical Design Considerations.—Much of the performance capability of coaxial-line oscillators is determined by the initial design.

It is impossible with tuning and loading adjustments to compensate for poor choice of line dimensions, poor feedback, and circuit elements having low unloaded Q . Generally, the circuit designer is faced with a set of performance specifications that include, electrically, such things as frequency range, minimum acceptable output power, and modulation characteristics. Physically, the complete unit may have to satisfy stringent requirements as to over-all weight, length, accessibility of parts, and ease of tuning and may have to operate under severe conditions of heat, cold, vibration, and altitude. The relaxation of one or more of these requirements normally leads to a possible simplification of the problem but, as a general case and for explanatory purposes, a circuit design will be outlined in Secs. 15-16 to 15-25 that attempts to satisfy most of the specifications.

15-16. Tube and Circuit Selection.—It is assumed that a tube is available that will deliver the power output required at the maximum operating frequency to be considered. Because of the unsuitability of multiple-tube circuits to wide-range operation, a single-tube unit is generally preferred.¹ For example, a type 2C38 will satisfy the operating-range and power characteristics outlined in Figs. 15-5 and 15-6. The manufacturer's tube characteristic data will indicate power-supply requirements, will include information on interelectrode capacitances, and will list the maximum allowable current, power-input and power-output ratings.²

Ultrahigh-frequency tubes are small in relation to their power-handling ability. This is the result, in part, of the use of efficient radiators for forced-air cooling but requires that provision be included for this air cooling. Typical radiator efficiencies are such that a blast of about 15 cu ft per min will adequately limit seal temperature in a small triode dissipating 100 watts at the plate. Tubes designed to work at higher power levels are often water-cooled.

In order to expose the plate radiator to the air stream, the basic circuit layout discussed will be the concentric-cavity type of oscillator shown in Fig. 15-15, in which the plate-grid line is concentric with the grid-cathode line. Several additional advantages accrue through the use of this configuration. The over-all length is materially reduced over the double-ended circuit of Fig. 15-4a, the tube is readily accessible for replacement purposes, and the positions of both plungers can be con-

¹ A multitube circuit is discussed in Sec. 15-28.

² While many types of tubes, such as those shown in Fig. 15-2, have been designed for coaxial-circuit use, some of these are still in the experimental stage. It is therefore impossible to list final physical and electrical characteristics. However, as an indication of performance, data for several early experimental types are listed in Table 14-1.

trolled from one end through an appropriate tuning mechanism such as push rods or lead screws.

15-17. Choice of Line Characteristic Impedances.—The choice of concentric-cylinder diameters must be based on mechanical as well as electrical considerations. The diameter of the innermost line, 8, in Fig. 15-15, is normally chosen as small as possible in order to permit a reasonable over-all diameter for the unit. Its minimum value is limited by the size of the vacuum-tube cathode and filament connector. It may need to be sufficiently large to carry, internally, such things as filament leads

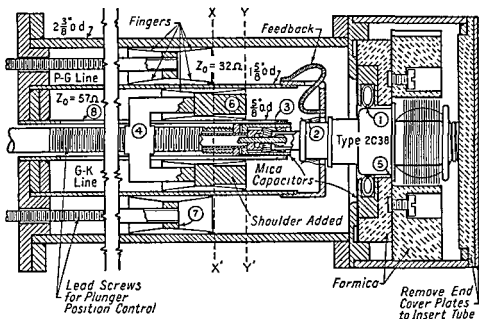


FIG. 15-15—Concentric-cavity type of coaxial-line oscillator. (1) Plate contact; (2) grid contact; (3) cathode contact, (4) nut and yoke driving grid-cathode plunger through slots in the cathode line; (5) tube stop properly to position tube in oscillator; (6) grid-cathode plunger; (7) grid-plate plunger; (8) slot in cathode line.

or a lead screw to drive the grid-cathode-line plunger through a slot in the line, or both. It should have a circumference large enough to provide a large contact surface between the cylinder and the movable plunger and, perhaps, to permit the incorporation of a bias-blocking capacitor. These considerations will probably limit the minimum diameter to at least $\frac{1}{2}$ in.

The diameter of the outer cylinder is usually kept small in the interests of physical size and the suppression of interfering modes. Present tubes are capable of operating at frequencies so high that this latter requirement becomes practically impossible to meet, and the problem of the suppression of interfering resonance modes must be solved through the avoidance of their propagation. This subject is discussed in Sec. 15-8.

The outer tube diameter must be sufficiently large to include practical short-circuiting plungers and tuning mechanisms in both concentric circuits. The choice of middle cylinder diameter must be consistent with the tube grid-terminal diameter and must result in line characteristic impedances meeting tank-circuit Q requirements (Sec. 15-7) and tuning-mode separation requirements (Sec. 15-6). Physical characteristics of air-filled coaxial-tuning lines are indicated in Fig. 15-1, and a typical set of diameters and characteristic impedances is shown in Fig. 15-15.

With these line-impedance values and with plate-grid and grid-cathode interelectrode capacitances of about 2 and 6.5 μmf , respectively, corresponding to the 2C38 tube, the ratio of the products of characteristic impedances and terminating capacitances is about 6 [see Eq. (15-2)], a sufficiently high ratio to ensure good mode separation.

If the unit were to be operated only at frequencies attainable in the $\lambda/4$ mode of operation, perhaps below 1000 Mc, it would be advantageous to increase the characteristic impedance of the plate line above the present 32 ohms to obtain a lower operating Q and wider modulated band widths. In this particular case, however, the unit must also tune from 1000 to 2500 Mc in the $3\lambda/4$ mode and the tank-circuit characteristic impedance of 32 ohms is considered satisfactory.

15-18. Physical Length of Lines.—Over-all line lengths are primarily determined by minimum-frequency requirements. From Eq. 15-5 and a knowledge of tube capacitance terminating the lines, it is possible to determine a maximum electrical length θ_1 such that, with the characteristic impedances as chosen in the previous section, the terminating capacitance will be resonated at the minimum desired frequency. From θ_1 , the physical length can be found from Eq. (15-7). For resonance

$$X_{cgp} = Z_{01} \tan \theta_1 \quad (15-5)$$

$$X_{cok} = Z_{02} \tan \theta_2 \quad (15-6)$$

$$l = \frac{0.83\theta}{f} \quad (15-7)$$

where θ = electrical line length, degrees

f = frequency, megacycles per second

l = length, meters

Z_{01} = characteristic impedance of the grid-plate line, ohms

Z_{02} = characteristic impedance of the grid-cathode line, ohms

X_{cgp} = reactance of the capacitance terminating the grid-plate line, ohms

X_{cok} = reactance of the capacitance terminating the grid-cathode line, ohms

The capacitive reactances of the terminating capacitances are those at the minimum frequency, and θ is the electrical line length at this frequency.

The calculated line length, such as l_1 in Fig. 15-7a, should be increased perhaps 10 per cent to take care of the approximations involved and the variations in tube capacitance and should be added to the axial length of the shorting plunger to determine the over-all required length.¹

This calculation has been based on the minimum operating frequency in the $\lambda/4$ mode. The maximum operating frequency in this mode can be estimated from a knowledge of physical line length with the plunger moved to its maximum forward position, nearest the tube (usually limited by feedback or output-coupling elements). This upper frequency cannot be found directly from Eqs. (15-5), (15-6), and (15-7) because it appears both in the argument of a transcendental function and outside the argument. That is

$$\frac{1}{2\pi fC} = Z_0 \tan \frac{2\pi l f}{v} \quad (15-8)$$

Since in the practical case, at the upper end of the mode, the electrical length is small enough to allow the replacement of $\tan \theta$ by θ , the following approximate formula can be used for this upper frequency:

$$f = 2760 \frac{1}{(Z_0 l C)^{1/2}} \quad (15-9)$$

where f = frequency, megacycles

Z_0 = characteristic impedance, ohms

C = terminating capacitance, micromicrofarads

l = line length, meters

Because of the increased influence of end effects for the now short sections of lines, the approximation is only fair. Values of f as given by Eq. (15-9) are too high by perhaps as much as 20 per cent. Since there must be no discontinuity in frequency coverage, it is important that at the upper frequency limit of the $\lambda/4$ mode there be sufficient line length (l_1 in Fig. 15-7c) behind the plunger to permit resonance of the line on the $3\lambda/4$ mode at this frequency with the plunger moved back into the cavity. If this is not true, the over-all physical length of the line must be increased to cover this contingency.

¹ As an example, if the oscillator in question must tune to 300 Mc, θ_1 can be found from Eq. (15-5) to be

$$\theta = \tan^{-1} \frac{X_{c22}}{Z_{01}} = \tan^{-1} \frac{265}{32} = 83.1 \text{ deg}$$

and l is found from Eq. (15-7) to be

$$l = \frac{0.83 \times 83.1}{300} = 0.23 \text{ meters}$$

so that, with allowances for plunger length, the over-all barrel length should be about 11 in.

Silver-plated brass tubes are generally used for the concentric cylinders. Wall thicknesses are commonly $\frac{1}{32}$ to $\frac{1}{8}$ in., the heavier tubing supplying the mechanical rigidity needed for the outermost cylinder.

15-19. Circuit Tuning Mechanisms.—Coaxial-line tuning is accomplished normally by a variation in the electrical length of line through a change in the position of an effective short circuit. This is accomplished through variation in the physical position of a short-circuiting plunger, such as shown in Fig. 15-1. Less frequently used, because of tuning-range limitations, are the methods illustrated in Fig. 15-16.

Plungers constructed for wide-frequency-range use are usually one of two types. When noiseless rapid tuning is required, it is possible to construct short-circuiting plungers of the resonant-choke type which do not contact the cylinder walls. These are discussed in detail in Sec. 32-29. Their frictionless, noiseless action is particularly adapted to sweeping local oscillators of wide-range receivers. A tuning range of at least three octaves may be obtained by proper design. At frequencies below 1000 Mc, the resonant-choke plungers tend to become excessively long. At all frequencies, they have a disadvantage that, to avoid propagation of interfering coaxial modes, they must be concentric with cavity walls to a tolerance so close that it is difficult to attain in physically large units. Wide-range resonant-choke plungers require particularly close plunger-to-wall clearances in addition to the concentricity requirements.

For transmitter units covering very wide frequency bands, the most successful plungers are those in which an actual d-c short circuit to the cylinder walls is made through sets of flexible fingers. An alternative contact is a coil-spring mechanism, one type of which is described in Sec. 15-20. The flexible fingers are most easily fabricated by slotting strips of sheet metal, usually 6-to-8-mil phosphor-bronze or beryllium-copper, which are then soldered to brass slugs forming the plunger base. Finger width is kept down to between $\frac{1}{16}$ and $\frac{1}{8}$ in. in order to increase

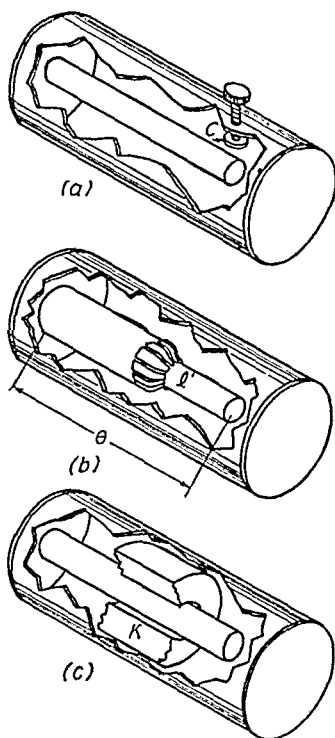


FIG. 15-16.—Tuning arrangements for coaxial lines: (a) line end-loaded by adjustable capacitance C ; (b) line length changed by sliding section L' ; (c) effective line length changed by variation in position of slug K of high dielectric constant.

the number of points of contact per plunger and to make the fingers flexible without being fragile. The best material is beryllium-copper, but it has the disadvantage of requiring a hardening treatment after the fingers are formed and before they are soft-soldered or after they are hard-soldered to the plunger base. The over-all assembly must be silver-plated. The lengths of the fingers are usually at least $\frac{1}{2}$ in. to provide flexibility. A finger plunger assembled in a coaxial-line circuit may be seen in Fig. 15-1. These fingers have been tipped with silver wire to improve contact.

The upper frequency limit of a circuit, when determined by interference of plunger fingers with feedback or coupling elements, can sometimes be increased by moving the position of the electrical short circuit forward through the addition of a shoulder on the plunger face. This is illustrated in Fig. 15-15 by the shoulder section that advances the short-circuit plane in the grid-cathode line from approximately XX' to YY' . Even this small addition will materially affect the resonant frequency in operation with short line lengths. Double sets of fingers are used to keep the plunger properly positioned in the cavity and to double the number of contact points as an aid in eliminating coupling of energy into the cavity back of the plunger.

15-20. Circuit-to-tube Connecting Systems.—A matter of concern in all coaxial-line circuits is the mechanical system for connecting the tube to circuit elements. This should have a structure permitting ready removal and insertion of tubes (plug-in action) and must make a multi-point contact around the entire circumference of the connector. The contact should be sufficiently flexible to take care of physical tolerances in tubes and to hold under vibration. It must be sufficiently positive to handle the very high circulating currents in the circuit elements.

Three satisfactory types of connectors are shown in Fig. 15-17. The best contactor for a particular application will depend on the tube design. However, the sleeve and spring contacting arrangements shown in Fig. 15-17 are preferred from two standpoints: (1) a wiping contact is made, and (2) there is less restriction on longitudinal motion of the tube when it is inserted.

The process of fabricating the fingers to make contact with the tube rings is similar to that described for the plunger fingers. The coil-spring contact of Fig. 15-17 is constructed by winding a coil spring of beryllium-copper wire of perhaps 0.015-in. diameter wire and soldering the ends together to form a ring. This ring, in turn, is forced into a circumferential slot in a disk-shaped holder, the slot width being less than the diameter of the coil. The spring is held in the slot by rolling the slot edges very slightly. The inside diameter of the coil ring, when in the slot, is slightly less than the outside diameter of the tube to be contacted,

which is always inserted with a slight twisting action causing the spring to "lie down" in the slot. The resulting deformation of the coil in two dimensions is such that each turn makes contact with the tube and slot walls, providing a large number of positive multipath connections of very short length.

Tube insertion is facilitated by arranging all sets of contacting fingers to point in the direction of insertion, as there is then less chance that the tube will foul any fingers. It must be emphasized that all fingers should make positive contact. Often poor operating efficiency and, in some cases, a complete lack of oscillation can be traced to failure of several adjacent fingers in a plunger or tube connector to make contact.

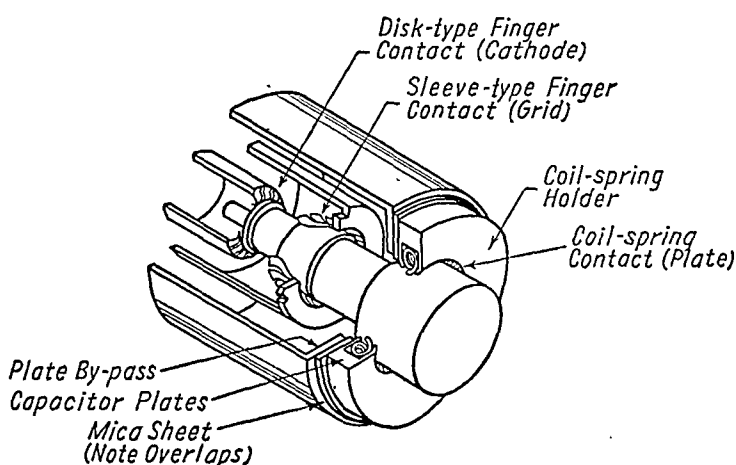


FIG. 15-17.—Tube contact arrangements.

There are several other methods of making contact especially suited to a particular tube type. These may vary from actually screwing a coaxial-line conductor to a threaded member in a tube, to an arrangement in which a tube-contacting terminal is bolted into the circuit through a split-flange mechanism. Neither makes for ready tube removal.

15-21. Feedback Design.—The ultimate performance of an oscillator will depend to a considerable degree on the type of feedback used. While the problem does not lend itself readily to exact analytical design, a general approach can be made following the discussion of Sec. 15-9. Cut-and-try alterations based on initial operating data usually make it possible to maximize performance.

Feedback requirements of the wide-range oscillator used for this general discussion can be met best by two interchangeable feedback systems, a low-frequency capacitive unit and a high-frequency inductive unit. There is generally considerable frequency overlap in the two, and a

convenient change-over frequency is often that at which the shift from $\lambda/4$ to $3\lambda/4$ mode operation is made. A typical arrangement is shown in Fig. 15-18. The low-frequency unit consists of three probes linking

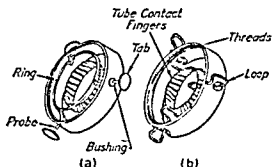


FIG. 15-18—Removable feedback units with tube-contacting fingers (a) capacitive feedback; (b) inductive feedback

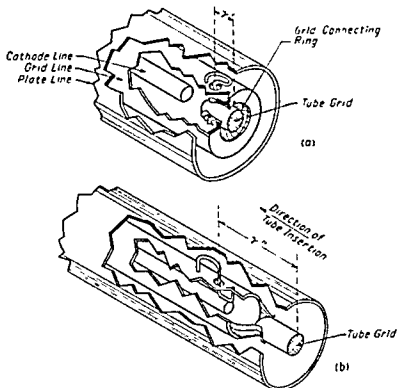


FIG. 15-19—Grid-terminal contacts and feedback location in a coaxial-line oscillator.
 $\gamma' \ll \gamma''$

the cavities through transformer bushings. Circular tabs terminate the probes in the tank cavity. Probes are joined by a wire ring in the grid-cathode cavity. Feedback is increased by increasing the size of tabs and

ring and their proximity to outer and inner conductors. The inductive-loop high-frequency unit consists of three S loops. The loop size in each cavity is kept small, feedback being increased by increasing the number of loops. One to perhaps six loops might be used, the choice of three offering a compromise. Each of these feedback elements is threaded to screw onto the end of the grid-line cylinder, and each carries a set of fingers to contact the tube grid connector. A cross-sectional view of a typical oscillator with the inductive feedback in place is shown in Fig. 15-15. Several types of feedback elements are shown in Fig. 15-12.

The grid-cylinder diameter of the cavity in Fig. 15-19a has been chosen larger than the grid connecting ring on the tube. This allows an advantageous location of both connecting fingers and feedback elements; the fingers point in the direction of tube insertion, and the plane of the feedback is advanced considerably toward the grid of the tube in Fig. 15-19a as compared with Fig. 15-19b.

Operational evidence of insufficient feedback is to be found in low grid current, inability to load the oscillator without stoppage of oscillations, and low operating efficiency. Excessive feedback is characterized by wide variations in the intensity of oscillation over a given frequency range, excessive grid current, broad grid-cathode-circuit tuning, and poor efficiency.

15-22. Direct-current Circuit Design.—Direct-current isolation of plate, grid, and cathode elements from each other is necessary in order to apply plate and bias voltage. There are two common methods for accomplishing this. In the first, the cylinders are isolated electrically from each other and the plungers have built-in capacitors. In the second, all cylinders operate at the same d-c potential, generally that of the grid, and isolating capacitors are built into the circuit at the tube. These methods are shown in Fig. 15-20. The scheme in Fig. 15-20b has a certain mechanical advantage in ruggedness in that the concentric cylinders can be bolted or soldered together for rigidity and for maintenance of concentricity.

By-pass capacitors are normally formed by mica-separated silver-plated brass plates. The plane of the capacitor is usually perpendicular to the longitudinal axis of the cavity, or the capacitor is constructed concentrically with a coaxial line. An example of the disk type is shown in Fig. 15-17, and a cylindrical type of capacitor is shown in Fig. 15-21. Both types are illustrated in Fig. 15-15. The disk type is easier to build, the cylindrical type can be built on a small-diameter line. The capacitance is increased by decreasing the thickness of the mica and by increasing the pressure on the plates. The use of silvered mica is convenient in increasing capacitance of small-area capacitors. For a given mica thick-

ness, two thin sheets are preferred to one thick sheet, as the use of two sheets tends to minimize breakdown through pinholes or cracks.

The effectiveness of the capacitors often cannot be judged by a low-frequency capacitance measurement. This is particularly true of the

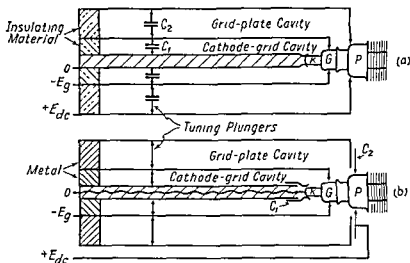


FIG. 15-20.—Direct-current circuit isolation in coaxial-line circuits. C_1 , bias blocking capacitor, C_2 , plate-voltage blocking capacitor.

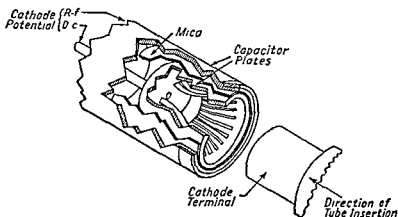


FIG. 15-21.—Cylindrical-type capacitor and finger contact.

cylinder type of Fig. 15-21, where the longitudinal length of the capacitor may be electrically great.

The capacitances required in bias and plate by-pass capacitors are dependent on several factors. The largest capacitance consistent with the physical space that can be devoted to the capacitor is usually desirable. This is tempered in the case of the bias-blocking capacitor by the need for reducing the grid-circuit time constant (Sec. 15-12) and by the

requirements of some modulation schemes. Typical values run from 200 to 2000 μf for oscillators operating in the 300- to 2500-Mc range.

High plate by-pass capacitance is also desirable to prevent low efficiency and *holes* in the frequency range, *i.e.*, regions of low or no output. If this capacitance appears in a wide-band modulation circuit, it may be necessary to limit its value considerably. As the desired tuning range is decreased, it is sometimes possible to set up the circuit with values of plate capacitance as low as 30 μf in single-ended circuits and still avoid "holes." For single-frequency operation, capacitances as low as 10 μf have been used. In wide-range oscillators, the figure probably will run nearer to 200 μf , and 500 μf would be desirable in an unmodulated oscillator.

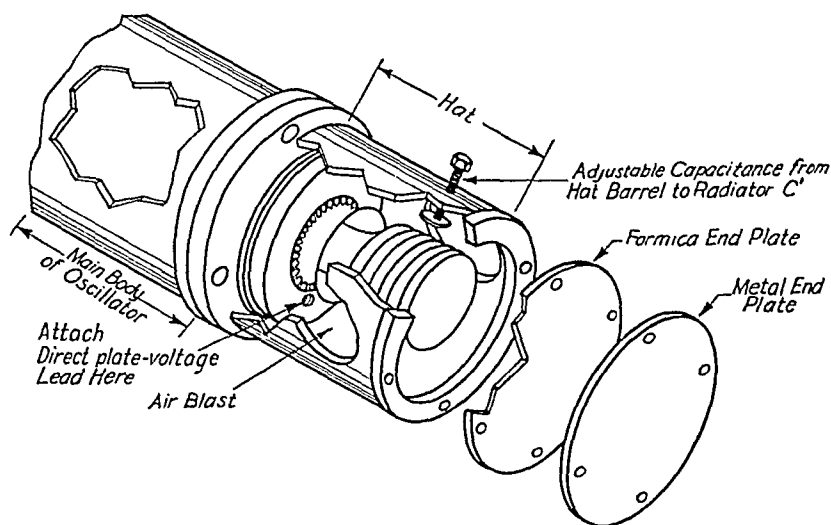


FIG. 15-22.—Hat design (air baffles not shown).

15-23. Plate Feed and Hats.—A cylindrical section, or *hat*, shown in Figs. 15-15 and 15-22 is often added to coaxial-line oscillators. It serves several purposes.

1. With an end-cover plate, it encloses the anode cooling radiator and serves to protect operating personnel from the high d-c voltage, between the anode and the main body of the unit.
2. It decreases r-f radiation from the cooling fins, which receive energy across the plate by-pass capacitance.
3. With appropriate baffles, it directs the air blast from the cooling fan onto the radiator.

Anode power is usually fed to the unit through a conductor entering the hat and tying to the anode side of the by-pass capacitance. A choke in the anode lead is generally unnecessary if it is possible to tie the lead to the by-pass capacitance plate, which is always at a low r-f impedance. This

is fortunate since it is practically impossible to construct a nonresonant choke covering perhaps a 10-to-1 frequency range.

15-24. Output Coupling.—Output power can be coupled from the tank circuit by an adjustable loop placed at a high-current point in the circuit or by a probe placed in a high-voltage field. Since the physical positions of high magnetic and high electric fields change with frequency, a considerable problem is encountered.

A loop placed in the face of the tank-circuit plunger is satisfactory but presents a difficult mechanical problem since it must move with the plunger, and its orientation must be adjustable for coupling variations.

A radially adjustable probe placed near the tube in the tank circuit is suitable for wide-range operation since the voltage field in this region is

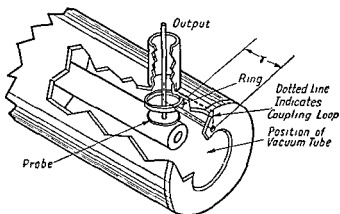


FIG. 15-23.—Probe and ring arrangement for wide-frequency-range coupling γ must be electrically short at highest frequency requiring loop coupling. Tube and tube terminal contacts are not shown.

reasonably constant and a probe can be made readily adjustable along a radius. If field conditions become unfavorable, two courses are open: (1) to move the probe a short distance longitudinally in the cavity to a point where more voltage exists, (2) to cause the probe to pass through a ring of such shape and so mounted in the cavity that when the probe tab is not near the ring capacitive pickup is possible, while with the tab moved up into the ring, the low capacitive reactance between ring and tab completes a loop suitable for inductive pickup, the inductive pickup being adjustable through the variable capacitance of tab to ring. This arrangement is shown in Fig. 15-23. The latter scheme has an advantage in that it can be made mechanically rugged and requires a coupling adjustment in a radial direction only.

A general discussion of coupling methods is given in Chap. 16.

15-25. Mechanical Design Notes.—Too much emphasis cannot be placed on mechanical details because they largely determine ultimate cir-

cuit performance. Finger and spring contacts must be positive. All portions of the circuit carrying r-f currents must be of high conductivity, preferably silver-plated.¹ Lossy material (Formica, etc.) should be kept out of the active part of the cavities. Butt joints that carry heavy currents (hat-to-cavity joint, for example) must be bolted securely or soldered. If bolted, the surfaces should be faced off before plating to assure flush contact, and a minimum of four and preferably six bolts should go into the bolt circle.

Sliding adjustments and contacts should be maintained under close tolerances, mechanically spring-loaded controls of coupling and tuning being incorporated where possible. As an alternative to push rods, tuning-plunger motion is controlled frequently by lead screws with a fine pitch. A gearing arrangement allowing either separate or simultaneous drive of cathode and plate plungers from a single dial often can be set up to achieve single-dial tuning control over incremental ranges sometimes as large as 2 to 1 or 3 to 1 in an oscillator covering a 10-to-1 frequency range.

A motion of a few thousandths of an inch in a tuning plunger or coupling probe has a marked effect on tuning or coupling at 2500 Mc, and a fine adjustment of controls must be possible.

15-26. Trouble Shooting and Holes.—It is the goal of wide-band oscillator design to set up a circuit capable of fully realizing the properties of a tube over very wide tuning ranges. Failure to obtain the desired performance is generally manifested in weak or intermittent oscillations, poor efficiency, and narrow frequency ranges in which no oscillations are present, *i.e.*, holes. Several typical troubles are outlined in this section.

Poor Over-all Efficiency.—At very high frequencies, tube capabilities may be a limitation. There may be poor tube or plunger-finger contacts. The feedback may not be correct. The unloaded cavity Q may be low through lack of silver plating, the presence of high-resistance joints, or the presence of lossy material in the cavity.

Broad Holes or Regions of Low Output.—The feedback may not be satisfactory in this range, the output coupling system may be inadequate, or an increase in the value of the by-pass capacitances may be needed.

Narrow Holes.—These regions, which may be only a few megacycles wide or which may be indicated only by a skip over a particular narrow frequency band, can usually be attributed to the coupling of an additional resonant circuit into the tank circuit, the breadth and intensity of the hole being dependent on the degree of coupling and the Q of the coupled

¹ Gold is sometimes used as a plating material. Also a very thin layer of rhodium or palladium is sometimes applied to a surface previously plated with a high-conductivity material as it affords a wear-resistant surface that withstands corrosion. The layer must be very thin to avoid a lowering of the conductivity.

circuit. The location of these resonant circuits can often be determined by exploration with a neon bulb. The following sources may be responsible for the trouble.

1. One source may be back-cavity resonance in which the section of the coaxial circuit behind the plunger resonates at the operating frequency (Fig. 15-1) and is coupled to the tank circuit through poor plunger-finger contact. This can usually be avoided through proper plunger construction with double sets of contact fingers.

2. Slot resonance may occur in which longitudinal slots in cavity walls, usually present to connect plungers to tuning mechanisms, couple a resistance into the tank circuit. Such slots should be avoided in the high- Q plate tank circuit but are usually permissible in the driving or grid-cathode circuit. The resonance can be broken up by short-circuiting the slot at several points along its axis.

3. Simultaneous resonance of the tank circuit in both the normal TEM coaxial mode and the $TE_{1,1}$ coaxial mode (see Sec. 15-8) may be responsible. Coupling to this undesired mode generally can be minimized by maintaining a symmetrical construction in the cavity and avoiding field distortions that would promote its propagation. A nonuniform circumferential distribution of by-pass capacitance sometimes causes this mode interference. Such interference will not be experienced when operation is in the $\lambda/4$ mode tuning range.

4. Another source of trouble is resonant circuits coupled through the plate by-pass capacitance. Even though the reactance, in ohms, of the plate by-pass capacitance may be very small, it is possible for this capacitance in combination with the impedance, at a particular frequency, of such things as the tube cooling radiator, plate lead, or plate choke to resonate and couple a resistance into the tank circuit. Trouble from plate leads and chokes can be avoided by the plate-lead connection outlined in Sec. 15-23 which eliminates the need for a choke.

The resonance of the cooling radiator as a resonant line in combination with the hat is more difficult to handle. Increasing the size of the plate by-pass capacitance will decrease coupling and eliminate the hole, but when modulation requirements make this impossible, it is necessary to resort to other means. Two ways are open. The Q of the resonant circuit in the hat can be lowered by the deliberate introduction of a lossy element. If the resultant Q is of the right magnitude, the effective resistance coupled into the main cavity at resonance can be lowered to the point where only a slight decrease in the intensity of oscillation will be noted.

This calls for a rather special Q value that may be difficult to set up, and the method is usually discarded in favor of moving the frequency of the hole by capacitance loading of the cooling-radiator resonant circuit. Thus, by the introduction of a variable capacitance in the hat, usually an adjustable tab, as C' in Fig. 15-22, the hole can be moved to one side of the desired operating frequency. The need for the tab capacitor C'

can be eliminated by a special construction of hat end plates, as described below.

By constructing the hat end cap in two sections, one Formica and one metal as in Fig. 15-22, it is sometimes possible to move the hole from one tuning range into another by simply removing the metal end plate, an expedient that serves the same purpose as C' . The oscillator shown operates in the $3\lambda/4$ mode range from 900 to 2700 Mc with the Formica and metal end plates in place. The capacitance of the metal plate to the radiator lowers the interfering resonant frequency of the hat cavity to 800 Mc. In the $\lambda/4$ mode range, 300 to 1000 Mc, the metal end plate is removed at the time feedback is changed from inductive to capacitive, and this removal of the end-loading capacitance changes the hole frequency from the low-frequency range to 1100 Mc. (In the low-frequency $\lambda/4$ band, there is little loss in radiation through the single Formica cover plate.)

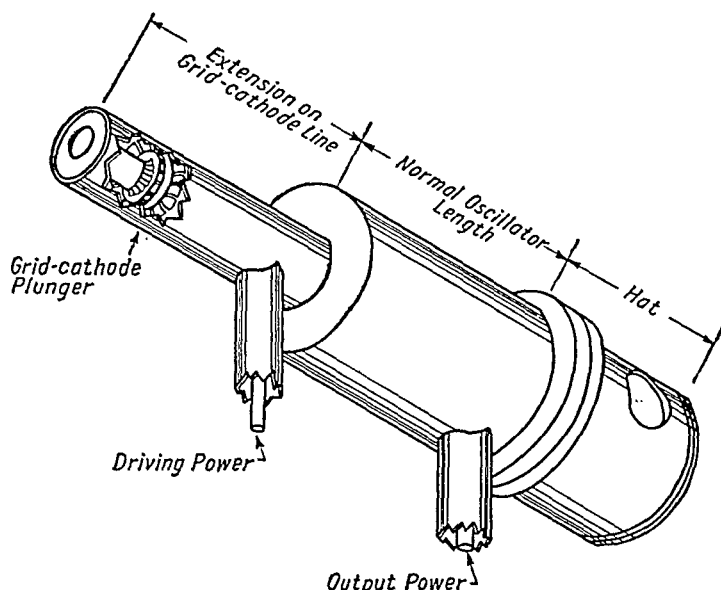


FIG. 15-24.—Amplifier unit with extended grid-cathode line to permit introduction of driving power into the internal grid-cathode line.

15-27. Modification for Amplifier Use.—While the discussion thus far has emphasized wide-range oscillators, most of the remarks apply equally well to coaxial-amplifier construction. For the most part, the unneutralized common-grid amplifier is physically similar to the oscillator circuits described except that no feedback elements are included and circuit isolation is maintained at a maximum.

A problem unique to amplifier operation in circuits of the type described here is that of the introduction of driving energy into the grid-

cathode circuit, which may be inside and concentric with the tank circuit. Two solutions are possible: (1) to drive the grid-cathode circuit through a loop in the plunger face, or (2) to expose a section of the grid-cathode line by increasing its length, as shown in Fig. 15-24, and then to operate on higher tuning modes in this line so that driving power may be coupled into the circuit forward of the plunger in any convenient manner. Performance of a type 3C22 triode in such a circuit is outlined at the end of Sec. 15-14.

By tuning the output circuit to the second harmonic of the driving-voltage frequency, the unit can be used as a frequency doubler. When a 3C22 doubler was driven with about 20 watts of driving power at a frequency of 650 Mc, the output was 5.2 watts at 1300 Mc with a plate efficiency of 4 per cent. The efficiency obtained by using the 3C22 as a Class B doubler does not compare favorably with the 11 per cent efficiency that can be obtained when the tube is operated as a Class B amplifier or oscillator at the same frequency.

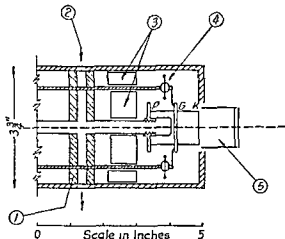


FIG. 15-25.—Liquid-cooled push-pull 3C22 oscillator. Opposite end identical to the end shown: (1) supporting rings of MF Formica; (2) cooling liquid, d-c plate voltage, and output coupling loop introduced through hole at zero voltage point; (3) noncontacting slug plungers; (4) feedback probe; (5) water-cooled type-3C22 (plate cooling fins removed).

15-28. Multitube Circuits.—A push-pull grid-separation type of coaxial-cavity oscillator that uses a pair of 3C22 lighthouse triodes is shown in Fig. 15-25. Liquid cooling was used in the model shown in order to enable the tubes to dissipate 200 to 300 watts. Power output in excess of 100 watts can be obtained from this oscillator over the frequency range from 500 to 600 Mc (about twice the single-tube output). Tuning is accomplished by moving the nonshort-circuiting plungers symmetrically from the center point of the cavity. Grid-cathode and grid-plate lines are electrically $\frac{1}{2}$ wavelength long, but capacitance-loading effects shorten the physical length considerably. Inasmuch as the center

point of the lines is at a voltage node, the liquid cooling lines can be brought out without disturbing the oscillator. The output coupling loop is also inserted at the high-current center point.

An obvious advantage of this type of oscillator construction is the low plate-to-ground capacitance (40 to 50 μmf), which is very desirable from the standpoint of wide-band plate modulation. If wide-band plate modulation is to be used on such an oscillator, it is necessary to use low-capacitance filament transformers and to modulate across their capacitance and the stray capacitances in the oscillator itself.

The minimum diameter of the innermost concentric tube in the unit, shown in Fig. 15-25, normally would be limited to excessively large values by the diameter of the 3C22 cathode terminal or plate-cooling fins. This trouble was partly circumvented by unsoldering and removing the cooling fins and inserting into the anode post a threaded member that could be incorporated in a liquid plate-cooling system. This alteration presents a considerable problem in the avoidance of tube fracture, and the problem could be more easily met with a unit using tubes such as the 2C38 with the grid-cathode circuit included as the innermost cavity in order to expose the plate radiators for forced-air cooling.

A second type of multitube coaxial-circuit oscillator is the so-called *shotgun* oscillator, an adaptation of concentric and parallel-line circuits, which was designed to make possible a push-pull oscillator with a wide tuning range. Two type 2C44 lighthouse triodes were used in a grounded-grid circuit, the physical configuration of which is similar to that shown in Fig. 17-25. Plate-circuit tuning is accomplished by a plunger that varies the length of the plate-grid cavity. The cathodes are tuned by plungers in two concentric lines that form the grid-cathode tuned circuits. A tuning range from 350 Mc to 840 Mc has been obtained with this arrangement in operation on the fundamental or $\lambda/4$ mode. Efficiencies of about 15 to 20 per cent were obtained at 600 Mc, with simple capacitance feedback consisting of rings in each of the cathode-grid cavities connected to a large ring in the plate-grid cavity. (This type of feedback is not shown in Fig. 17-25.)

CHAPTER 16

POWER-OUTPUT COUPLING METHODS

BY J. G. STEPHENSON

16-1. Output Coupling to Antenna or Load.—Devices for coupling to an antenna or load form an extremely important part of any u-h-f oscillator or amplifier, as the degree of load coupling largely determines the effective value of shunt resistance R across the tank circuit. The desired magnitude of this resistance is fairly well fixed by the voltage and current relations required in Class C oscillators or amplifiers for efficient performance. If the coupling is too "tight," R becomes low, the efficiency is reduced, and, in oscillators, oscillation may cease altogether. The ability to stop oscillation by adjusting the coupling is a good criterion that the maximum degree of coupling is at least adequate. If the coupling is too "loose," the shunt resistance is so high that little or no power is delivered to the load.

Optimum coupling can best be determined with the aid of some form of r-f output voltage indicator. When this is not available, a simple rule of thumb, which has proved applicable to a wide variety of u-h-f oscillators, is to increase the load coupling until the oscillator grid current decreases to approximately half of its unloaded value. Under these conditions the output will usually be close to optimum.

A device for coupling the tank circuits of u-h-f oscillators or amplifiers to a dissipative load through a transmission line must satisfy a number of requirements:

1. The coupling device must act as a transformer to match the low impedance of the transmission line to the relatively high impedance of the tank circuit.

2. The coupling system must enable operation with a fairly high standing-wave ratio in the outgoing transmission line. This line may be terminated in its characteristic impedance Z_0 , in which case the standing-wave ratio is unity and the device matches a resistive impedance. More often in practical cases, however, the standing-wave ratio is greater than unity and may be as high as 4:1. This requires that the device operate with a load impedance having appreciable reactance.

3. The device may have appreciable internal reactance, which must be tuned out to obtain proper impedance match.

4. The coupler must have low internal losses.

5. In most cases, the output coupler must work into an unbalanced coaxial line, although it is sometimes desirable to operate into a balanced two-wire transmission line.

6. For coupling to push-pull circuits, the output coupler must load the circuit symmetrically, and its presence cannot be allowed to introduce serious electrostatic or magnetic unbalance into the tank circuit.

7. It is very desirable in wide-range oscillators that the degree of coupling be as nearly constant as possible as the frequency changes. This avoids the necessity for constant readjustment of the coupling.

8. When adjustment of coupling is necessary, it should be mechanically convenient, and positive electrical contact to moving parts must be maintained.

16-2. Effect of Standing-wave Ratio on the Output of Ultrahigh-frequency Transmitters.—Tests on typical u-h-f oscillators and amplifiers

to determine the effect of standing-wave ratio in the output transmission line on the power obtainable show results similar to those of Fig. 16-1. This shows the reduction in power output as the standing-wave ratio increases for a typical balun-type output coupler and a Class AB r-f amplifier at about 200 Mc. Similar results are observed for other coupling methods and at other frequencies, with variations of perhaps ± 15 per cent from the values shown in Fig. 16-1. In

no case did the average reduction in power exceed 30 per cent up to a standing-wave ratio of 2.5:1.

The data were taken by adjusting first for the desired standing-wave ratio with a parallel stub in the outgoing transmission line, and then moving the standing wave with respect to the transmitter by moving the stub along the line, meanwhile keeping the transmitter adjusted for maximum output. The point of minimum output obtainable with a given standing-wave ratio was then recorded. It represents the worst condition that can exist for the transmitter at the particular standing-wave ratio.

16-3. Untuned Coupling Loop.—Untuned coupling loops are frequently used in coupling into cavities, wavemeters, magnetrons, etc. In such applications they are ordinarily fixed in size and position so that the degree of coupling is also fixed. Wide-range oscillators require that the coupling can be varied, and therefore in such applications untuned

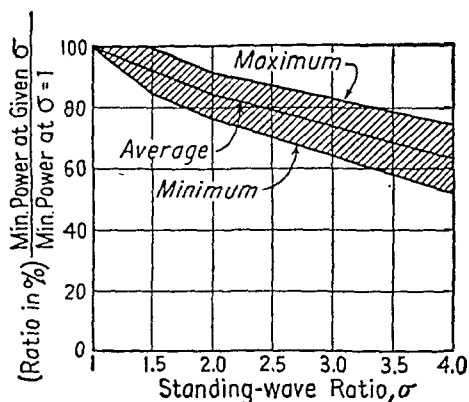


FIG. 16-1.—Effect of standing-wave ratio on power output of transmitter.

loops are usually constructed so that either the position or the orientation of the loop with respect to the magnetic flux lines, or both, can be conveniently varied. A typical example of the untuned loop is shown in Fig. 16-2, where the loop is located near the face of the tuning plunger in a coaxial-line oscillator. It may be turned so as to link more or fewer flux lines and thus vary the coupling. Since it is at a current-maximum point in the line for all frequencies, its axial position with respect to the plunger does not need to be varied.

From the equivalent circuit of Figs. 16-2b and c, it is apparent that the induced voltage E must be great enough to develop the required power output across the load resistance R_L through the series reactance of the loop (ωL_1) and the load reactance X_L . The induced voltage E is proportional to the flux density, the frequency, the area of the loop,

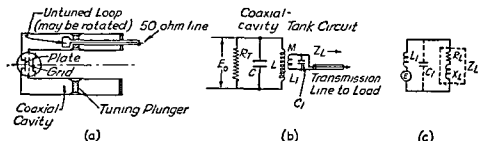


FIG. 16-2—Untuned-loop coupling to a coaxial-oscillator tank circuit (a) untuned loop in a coaxial cavity, (b) schematic diagram of untuned loop coupled into cavity; (c) equivalent circuit of untuned loop. R_T , equivalent-loss resistance due to tube and circuit losses; L , C , parameters of tank circuit; M , mutual inductance between loop and cavity; L_1 , self-inductance of loop; Z_L , input impedance of 50-ohm transmission line terminated in load $= R_L \pm jX_L$; E_s , r-f voltage across tank circuit; E , induced voltage in loop.

and the cosine of the angle between the direction of the magnetic flux and the normal to the plane of the loop. For typical loops used in the u-h-f region the self-reactance is comparable in magnitude to the input impedance of the 50-ohm transmission line. Thus a greater induced voltage is required for untuned loops than for loops provided with some arrangement for "tuning out" the circuit reactance. In general, therefore, untuned loops must have a greater area than tuned loops to obtain the same degree of coupling and must be located at a point of high magnetic flux density in the output tank circuit. Such conditions can usually be met in cavities or coaxial-line tank circuits where the fields are closely confined and concentrated, but in the usual parallel-line tank circuit, for example, it is difficult to find a point where the magnetic flux density is high enough to permit adequate inductive coupling with an untuned loop of reasonable size. The self-inductance of loops can be minimized by using large-diameter wire.

When untuned loops have dimensions small compared with a wavelength, they are conveniently treated as an equivalent generator in series with the lumped inductive reactance ωL_1 , as in Fig. 16-2c. If the perimeter of the loop exceeds about 0.15λ , however, it may be desirable to consider the loop as a short transmission line with distributed constants. Coupling behavior can then be interpreted in terms of the standing waves on the loop.

Experience has shown that untuned coupling loops cannot readily cope with high standing-wave ratios in the outgoing transmission line. When the input impedance Z_L of the outgoing transmission line has a high reactive component, the only degree of flexibility available in the untuned loop is to increase the coupling as far as possible, but this is frequently inadequate. Some means of tuning the loop provides much greater adaptability to high standing-wave ratios.

The equivalent circuit of Fig. 16-2c shows a very small shunt capacitance C_1 across the loop. This has been neglected in the discussion up to this point because usually it is so small that no self-resonance effects are experienced with ordinary loops in the decimeter wavelength region. Such problems can become serious as the wavelength approaches 10 cm, however, and may also be encountered at longer wavelengths if an unusually large loop or one with several turns is used.

16-4. Loop Tuned with Series Stub.—Figure 16-3 shows a method of coupling which employs a tuning stub in series with the loop and the input of the transmission line. The stub provides additional flexibility in adjustment as well as permitting variations in the amount of coupling, and enables more effective coupling when there is a high standing-wave ratio in the outgoing transmission line. While this method has been used with coaxial tank circuits, it has found its widest application in parallel-line tank circuits where the magnetic flux density is seldom high enough to permit adequate coupling with an untuned loop (see Sec. 16-3). Even with the series tuning arrangement, it has been necessary to employ loops of fairly large dimensions (with perimeters of the order of 0.15 to 0.30λ) in order to obtain adequate coupling to parallel-wire lines over the tuning range desired. It is convenient to consider such a loop as a portion of transmission line of fairly high characteristic impedance (ordinarily several hundred ohms) and to study the coupling system in terms of standing waves rather than to treat it in terms of lumped parameters.

Inductive coupling is maximum when the loop is placed close to the parallel lines near the short-circuiting bar so that the loop can link the maximum magnetic flux. Practically, however, in the case of parallel-wire lines the situation is complicated by the fact that usually there is an appreciable amount of capacitive coupling present in addition

to the inductive coupling. The relative amounts of capacitive and inductive coupling vary widely in different cases; sometimes the coupling is

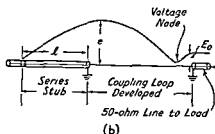
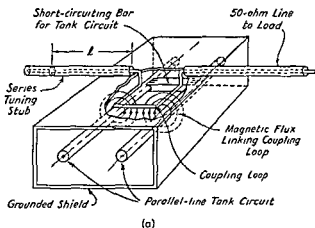


FIG 16-3.—Series-tuned coupling loop with parallel-line tank circuit: (a) physical construction; (b) typical standing wave of voltage in coupling system; e , voltage at any point along stub or coupling loop; E_0 , voltage across input to 50-ohm transmission line; L , length of tuning stub

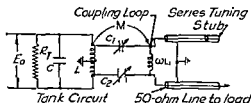


FIG. 16-4.—Schematic diagram and equivalent circuit of series-tuned loop coupled to a parallel-line tank circuit: R_r , equivalent loss resistance due to tube and circuit losses; M , mutual inductance between loop and tank circuit; C_1 , C_2 , equivalent coupling capacitances representing capacitive coupling to loop from both parallel lines (magnitudes of C_1 and C_2 not the same because of standing wave on loop); L , C , parameters of equivalent tank circuit; E_0 , r-f voltage across tank circuit; ωL_1 , self-reactance of loop (or self-impedance of loop considered as a short transmission line with ground return).

almost entirely inductive, sometimes principally capacitive. Available experimental evidence suggests that capacitive coupling becomes impor-

tant when (1) the loop is physically large, *i.e.*, the perimeter approaches or exceeds 0.25λ ; (2) the loop is made of a wide flat conductor; and (3) the loop is close to the parallel lines and near the high voltage point. The unequal voltage distribution on the loop due to the presence of a standing wave is taken into account in the schematic diagram of Fig. 16-4 by the capacitances C_1 and C_2 of unequal magnitude between the two sides of the loop and the tank circuit. There is some degree of unbalance in the capacitive coupling, but it is apparently not serious.

The series-tuned coupling loop, though mechanically simple, is electrically quite complicated and does not lend itself readily to quantitative analysis. The following characteristics, determined largely experimentally, summarize the performance of series-tuned loop coupling.

1. When considering purely inductive coupling, the series-tuned loop enables "tighter" coupling with a given loop area in a given magnetic flux density than is obtainable with an untuned loop. This is due to the impedance-matching properties of the tuning stub.

2. When coupled to a parallel-line tank circuit, the loop must usually be made large in terms of a wavelength (perimeter of 0.15 to 0.25λ) to obtain satisfactory coupling. As a consequence, standing waves of voltage appear on the loop.

3. When coupled to a parallel line, considerable capacitive coupling is frequently present in addition to the inductive coupling. The capacitive coupling may even predominate under certain conditions.

4. The length of the series stub is usually such as to place a voltage node on the loop at a point near the input to the transmission line which goes to the load (see Fig. 16-3). In typical cases the stub is 0.15λ to 0.24λ in length. (Exceptions to this may occur when the coupling is largely capacitive.)

5. The additional flexibility afforded by the series stub allows the coupling system to cope with almost any standing-wave ratio in the outgoing transmission line. However, when the standing-wave ratio is greater than about 2:1 or 3:1, the tuning of the stub becomes very critical.

6. The frequency band over which the coupling system will operate effectively without retuning varies widely in different cases. When the coupling is largely inductive, the tuning of the stub is fairly sharp and the equivalent Q of the output tuning circuit may be of the order of 10 with unity standing-wave ratio in the transmission line, and perhaps 20 or more with higher standing-wave ratios. When the coupling is predominantly capacitive, the stub setting is not at all critical, and the band width is correspondingly greater.

7. The use of a tunable stub makes it more difficult to design a compact coupling unit to cover a wide frequency range. However, the type shown in Fig. 16-5 tunes from approximately 200 to 500 Mc. Note the removable connector that permits the insertion of an additional length of line for tuning at the low-frequency end of the range. Figure 16-5 also shows an inductively loaded loop that enables a given coupling system to be resonant at a lower frequency than would otherwise be possible.

16-5. Balun or Balanced-unbalanced Transformer Type of Output Couplers.—Coupling systems based on the well-known *balun*, or balanced-to-unbalanced transformer, have proved very satisfactory in a variety of applications to u-h-f transmitters, particularly in the frequency range from 100 to about 700 Mc. Balun couplers consist essentially of two parallel coaxial lines, as shown in Fig. 16-6, with the center conductor emerging from the end of one line and tapping to the end of the outer conductor of the other. The outer conductors of the balun coupler

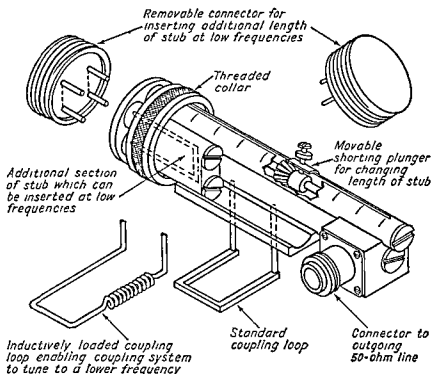


FIG. 16-5.—Series-tuned loop coupling unit. Frequency range 200 to 500 Mc.

are then inductively, conductively, or capacitively coupled symmetrically to a tank circuit instead of continuing into a two-wire transmission line as they would in the ordinary balanced-unbalanced transformer. Whenever they can be used, baluns are superior to tuned-loop coupling systems because they are much less critical in adjustment and can more easily be made to function over a wide frequency range. A detailed discussion of balun transformers is given in Secs. 3-13 to 3-18.

Some advantages of this type of coupler are as follows:

1. Almost complete electrostatic symmetry is provided, making it an ideal method for coupling to push-pull tank circuits without causing unbalance.
2. Balun couplers convert from balanced loading of the tank circuit to an unbalanced coaxial line, which is the type of output line ordinarily used. Two-

wire balanced lines can also be handled by this type of coupler with only a minor modification.

3. Most balun-type couplers function satisfactorily over a wide frequency band (1.5-to-1 or 2-to-1) without any adjustment, *i.e.*, the coupling does not vary rapidly with frequency. This is a big advantage in wide-range transmitters. If the balun short-circuiting bar is adjustable, such a coupling unit can be designed to accommodate a very wide frequency range (at least 5-to-1).

16-6. Inductively Coupled Balun.—Figures 16-6*a* and *b* show an inductively coupled balun with a parallel-line tank circuit. The equivalent

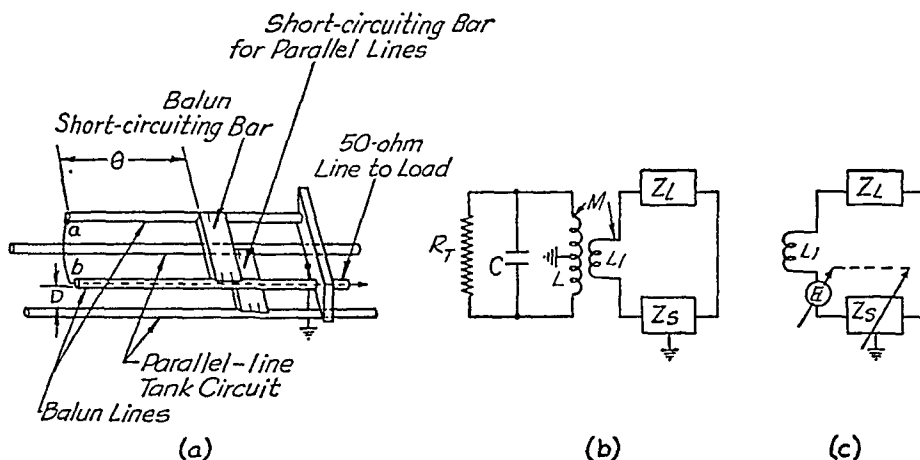


FIG. 16-6.—Inductively coupled balun and equivalent circuit: (a) physical construction of inductively coupled balun; (b) schematic diagram of balun inductively coupled to parallel-line tank circuit; (c) equivalent output circuit of inductively coupled balun; θ , electrical length of balun line; D , separation between balun line and parallel lines of tank circuit; R_T , equivalent loss resistance due to tube and circuit losses; L , C , parameters of tank circuit; L_1 , self-inductance of wire across end of balun; Z_0 , characteristic impedance of balun line; Z_s , impedance between points *a* and *b* of balun line $= jZ_0 \tan \theta$; Z_L , input impedance of 50-ohm line terminated in the load; E , voltage induced in balun loop; M , mutual inductance between balun loop and tank circuit.

lent circuit may be represented as in Fig. 16-6*c*. The voltage E is induced by magnetic flux linkages in the loop formed by the balun lines and the terminating wire across the end of the balun. This induced voltage is proportional to the area of the loop, the flux density linking the loop, and frequency. Z_s represents the impedance of the balun lines at the terminals *a* and *b* in Fig. 16-6*a* and may be expressed as

$$Z_s = jZ_0 \tan \theta \quad (16-1)$$

where θ is the electrical length, and Z_0 the characteristic impedance of the balun line. Z_s is effectively in series with the input impedance Z_L to the 50-ohm transmission line, which is terminated in the load. The outer conductor of the 50-ohm output line is also one of the two parallel lines forming Z_s . With the mid-point of Z_s grounded, point *b* in Fig.

16-6a is therefore above ground potential by the voltage across half the impedance of Z_s .

For a fixed position of the balun short-circuiting bar, the induced voltage E tends to increase with the frequency. The series impedance Z_s also increases, however, so that the resulting voltage across the input impedance Z_L of the transmission line tends to remain more nearly constant with changing frequency. This accounts partly for the broad-band characteristics of the inductively coupled balun.

The impedance Z_s , at the terminals of the balun line between the points a and b in Fig. 16-6a, is usually an inductive reactance, and its magnitude depends on the magnitude of the induced voltage E and, hence, on the magnetic flux density and the separation D of the balun

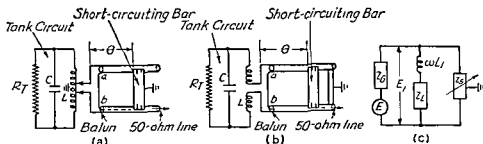


FIG. 16-7.—Conductively coupled balun and equivalent circuit: (a) balun tapped onto tank circuit; (b) balun incorporated into tank circuit; (c) equivalent circuit, Z_G , equivalent internal impedance of generator; E , equivalent generator voltage; Z_s , impedance between points a and b of balun line $= jZ_0 \tan \theta$; R_T , equivalent loss resistance due to tube and circuit losses; L , C , parameters of tank circuit; L_1 , self-inductance of wire across end of balun; θ , electrical length of balun line; Z_0 , characteristic impedance of balun line, E_1 , voltage appearing between taps on tank circuit; Z_L , input impedance to 50-ohm line terminated in the load.

from the parallel-line tank circuit. By adjusting either the spacing D or the length θ of the balun coupler, or both, a satisfactory degree of coupling can usually be obtained. For wide-frequency coverage it is convenient in some cases to gang the balun short-circuiting bar with the parallel-line tuning control so that the length of the balun decreases as the frequency increases, thus tending to maintain the desired amount of coupling. The adjustment of this, or any balun coupler, involves no resonant circuits and hence is not critical.

16-7. Conductively Coupled Balun.—The operation of the conductively coupled balun is similar to that of the inductively coupled type except that instead of a voltage being induced by a varying magnetic flux linking the balun loop, the ends of the outer conductors of the balun line are tapped directly across a portion of the tank circuit as in Fig. 16-7a. A variation is to employ the balun line as part of the resonant circuit, as shown in Fig. 16-7b. The method is applicable either to

parallel lines or coil tank circuits, but does not lend itself readily for use with coaxial lines.

In the equivalent circuit of Fig. 16-7c, the voltage E_1 appearing between the taps on the tank circuit is represented by an equivalent generator voltage E in series with an internal impedance Z_g . The voltage E , and to some extent the impedance Z_g , can be varied by changing the taps on the tank circuit. The voltage E_1 appears across Z_s and Z_L in parallel. Z_s is the impedance across the terminals of the balun line, and Z_L is the input impedance of the 50-ohm line which is terminated in the load. The degree of coupling to the tank circuit is increased by lengthening the balun line or by increasing the separation between the taps. Since the equivalent voltage E does not vary rapidly with frequency and no resonant impedances are involved in the circuit, the coupling remains reasonably constant over a fairly wide frequency band.

16-8. Capacitively Coupled Balun.—The capacitively coupled balun is very similar to the conductively coupled type with the difference that small capacitors are connected in series with the taps on the tank circuit, usually for the purpose of isolating the antenna circuit from the d-c plate voltage. The reactance of these capacitors is usually less than 50 ohms. The introduction of the capacitances in series with the balun merely requires a different adjustment of the taps or of the impedance Z_s to obtain the same degree of coupling as in the conductively coupled system.

16-9. Capacitance-probe Coupling.—Figure 16-8a shows capacitance-probe coupling as used in a coaxial-cavity oscillator. The probe is represented in the equivalent circuit of Fig. 16-8b as a small variable capacitance C_1 in series with a relatively high-voltage generator (the r-f tank-circuit voltage) and the input to the 50-ohm transmission line which is terminated in the load. The coupling is varied by moving the probe nearer to, or farther from, the center conductor of the coaxial line, thus varying the capacitance C_1 . In a typical coaxial oscillator delivering a power output of 40 watts into a 50-ohm resistive load at 500 Mc, the probe capacitance is about $0.6 \mu\text{mf}$, and the voltage across the input to the transmission line is of the order of one-tenth the r-f voltage in the coaxial line at the point where the probe is inserted.

The probe must be located at a point along the coaxial cavity where the electric field intensity (and hence the r-f voltage) is fairly high. For $\lambda/4$ mode operation the probe should be close to the tube end of the coaxial line. In wide-range oscillators utilizing both $\lambda/4$ and $3\lambda/4$ mode operation, it may happen at some frequencies in the $3\lambda/4$ mode that the probe falls at a voltage node in the coaxial line. Coupling can then be restored by moving the probe axially along the line to a point of sufficient electric-field intensity. Combined capacitive and

inductive coupling may also be used to overcome this difficulty (see Secs. 15-24 and 16-10).

The area of the capacitance tab on the probe should be great enough so that the desired coupling can be obtained with a reasonable spacing between the tab and the inner conductor of the coaxial line. Otherwise a voltage breakdown of air will probably occur at high-altitude conditions owing to the high r-f voltage existing between the tab and the inner conductor of the coaxial line.

For a fixed position of the probe, the degree of coupling does not change rapidly with frequency because no resonant phenomenon is

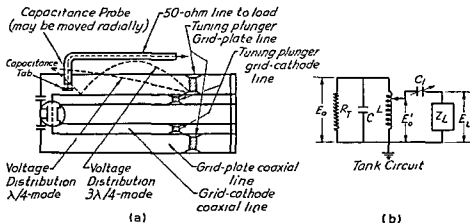


FIG. 16-8.—Capacitance-probe coupling into coaxial cavity: (a) physical construction and voltage distribution (1) $\lambda/4$ mode, (2) $3\lambda/4$ mode; (b) equivalent circuit; R_T , equivalent loss resistance due to tube and circuit losses; L , C , parameters of resonant tank circuit; C_1 , capacitance of probe to inner conductor of grid-plate coaxial line; E_L , voltage across input to 50-ohm line; Z_L , input impedance to 50-ohm line which is terminated in the load; E_0' , voltage across coaxial line at axial point where probe is inserted, E_0 , voltage across coaxial line at tube end.

involved. Adjustment of the coupling is required more frequently in operating into a line with a high standing-wave ratio, however, since the only way of compensating for a large reactive component of the input impedance of the transmission line is simply to increase the coupling until the desired loading is obtained. By making possible such adjustments, probe coupling is able to cope satisfactorily with fairly high standing-wave ratios.

The advantages of probe coupling may be summarized as follows:

1. Relative freedom from frequency sensitivity (as much as a 1.3-to-1 range in frequency may be covered under some conditions without readjustment of the coupling)
2. Ease and convenience of adjustment of the coupling with only one variable, i.e., the radial position of the probe

3. Adaptability to coaxial-type tank circuits, in which the electric field is intense and confined to a small volume

4. Ability to provide satisfactory coupling over a very wide frequency range in the coaxial type of tank circuit. Probe coupling is not so convenient with parallel-line tank circuits, for example, where the probe size must be considerably larger because of reduced electric field intensity.

16-10. Combined Probe and Loop Coupling.—At frequencies for which the standing-wave distribution along the axis of a coaxial tank circuit is such as to place the probe at a point of zero r-f potential, no capacitance-probe coupling can be obtained. By the addition of a

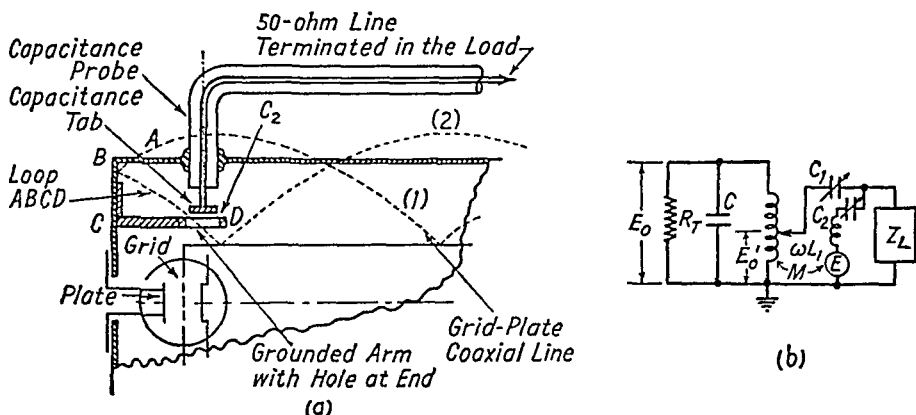


FIG. 16-9.—Application of combined capacitive and inductive coupling to a coaxial-line oscillator: (a) physical construction and distribution of (1) current, (2) voltage in grid-plate line; (b) equivalent circuit; C_1 , small variable capacitance between capacitance probe and inner conductor of coaxial line; E_0 , voltage across coaxial line at tube end; E , voltage induced in loop ABCD by varying magnetic flux; C_2 , small variable capacitance between probe and grounded arm which forms loop ABCD for inductive coupling; E_0' , voltage across coaxial line at axial position where probe is inserted; Z_L , input impedance to 50-ohm line going to load.

small bent arm with a hole at the end through which the probe can pass, it is possible to obtain inductive coupling. The probe can thus remain at the same axial location in the line. Figure 16-9a illustrates the construction. The arm forms part of a loop ABCD which is in series with the probe and the small capacitance between the tab and the arm. Thus the coupling can still be varied by moving the probe in or out. Except over a very small range of frequencies, the intensity of both the electric and the magnetic fields will be appreciable at the probe position, and therefore both capacitive and inductive coupling will be present with this coupling scheme. The equivalent circuit of Fig. 16-9b indicates both sources of coupling to the tank circuit (see Sec. 15-24).

16-11. Direct Coupling.—Direct coupling to a parallel-line circuit is shown in Fig. 16-10. Note that the tank circuit is grounded. This method has been used successfully with split-anode magnetrons, though

it is not ideal because of some unbalance that is introduced. The position of the tap T is varied until sufficient coupling is obtained. This occurs when the distance D is only a small portion of the total line length L , since the input impedance of the transmission line Z_L is small compared with the total impedance across the tube end of the parallel lines. The method is inherently unbalanced, as shown in the equivalent circuit of Fig. 16-10b, but has the advantages of simplicity and ease of adjustment, which may outweigh the disadvantage of unbalance. Ganging tap T with the short-circuiting bar of the parallel line will provide satisfactory coupling over a wide frequency range with no further adjustment. Practically, if the frequency range is more than about 2 to 1, some trimming of the tap position with respect to the short-

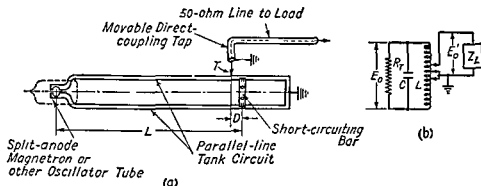


FIG. 16-10.—Direct coupling to parallel-line tank circuit. (a) physical construction; (b) equivalent circuit; R_T , equivalent loss resistance due to tube and circuit losses, L , C , parameters of tank circuit E_o , r-f voltage across tank circuit; E_o' , r-f voltage between shorting bar and tap, T ; Z_L , input impedance to 50-ohm line which is terminated in load; T , tap or point where output line is tapped to parallel line; D , distance between short-circuiting bar and tap T ; L , total length of parallel line.

circuiting bar may be necessary. This method is easily adapted to a balanced two-wire output line by simply tapping to each of the parallel lines.

One disadvantage of direct coupling is that certain parallel-line modes may be more heavily loaded than others, thus offering the possibility of mode jumping. Any mode of frequency such that a voltage node falls at or near the tap position will tend to be favored since it is lightly loaded, while the desired mode may be suppressed because it is more heavily loaded.

In coaxial circuits, direct coupling is often convenient, particularly for driving amplifiers, as illustrated in Fig. 15-25 of Sec. 15-27.

A method of obtaining balanced direct coupling into an unbalanced coaxial line is shown in Figs. 16-11a and b. This method differs from that of Fig. 16-10 in the use of a half-wave line connecting the two symmetrical coupling taps. This loads the tank circuit symmetrically

and provides a necessary phase reversal between one tap and ground, effectively paralleling the two sources of voltage. Actually, the loading will be truly balanced only at the frequency at which the transmission line connecting the two taps is electrically $\frac{1}{2}$ wavelength long.

The balun method illustrated in Fig. 16-7a is another direct-coupling scheme that permits unbalanced coaxial-line output with balanced loading of the tank circuit. It may be readily adapted to parallel-line circuits by tapping the two lines at a point ahead of the short-circuiting bar, as in Fig. 16-11.

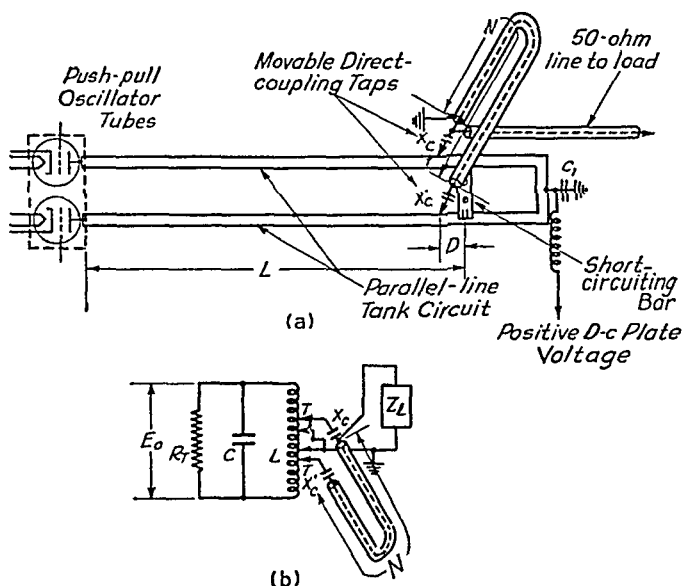


FIG. 16-11.—Balanced direct coupling to parallel-line tank circuit: (a) physical construction; (b) equivalent circuit; N , length of phasing line, $\lambda/2$ at center frequency of range; R_T , equivalent loss resistance due to tube and circuit losses; L , C , parameters of tank circuit; E_o , r-f voltage across tank circuit; Z_L , input impedance to 50-ohm line which is terminated in load; X_c , X_c' , reactances of series capacitors used for d-c isolation; C_1 , by-pass capacitance to ground; D , distance between short-circuiting bar and taps; L , total length of parallel line.

16-12. Series Output Coupler for Parallel-line Circuit.—This coupling method has the advantage of requiring no adjustments over approximately a 2-to-1 frequency range when the coupling system is properly designed. Furthermore, since no part of the coupling system extends beyond the short-circuiting bar, a higher maximum frequency may be attained than is obtainable with direct coupling. Figure 16-12a shows the construction, and Fig. 16-12b an approximate equivalent circuit. The parallel line is terminated by a short-circuiting bar as usual, except that one side of the line, instead of being short-circuited directly to ground, goes into a telescoping section that is effectively in series with the short-circuiting bar. The telescoping section is fol-

lowed by four impedance-matching sections in series, each $\frac{1}{4}$ wavelength long at the center of the frequency range. This arrangement effects a gradual impedance match from the input of the transmission line at point *P* (about 50 ohms) to the entrance of the telescoping section at point *Q* (10 to 15 ohms) over approximately a 2-to-1 frequency range. The telescoping section is of very low characteristic impedance (7 to 15 ohms) and is about $\frac{1}{4}$ wavelength long at the lowest frequency to be used, and $\frac{1}{2}$ wavelength at the highest frequency. The impedance in series with the short-circuiting bar at *R* will therefore be quite low (5 to 15 ohms) at all frequencies in the 2-to-1 tuning range because of the

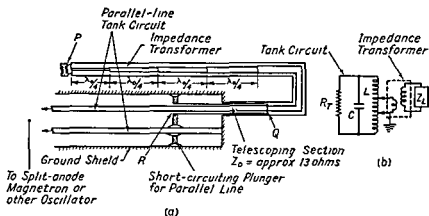


FIG. 16-12—Series output coupling system for parallel-line tank circuit: (a) physical construction; (b) equivalent circuit, R_T , equivalent loss resistance due to tube and circuit losses; L , C , parameters of tank circuit; Z_L , input impedance to 50-ohm line which is terminated in load; λ_0 , wavelength at center frequency of tuning range.

low characteristic impedance of the telescoping section. The relatively high current at the short-circuiting bar flowing through even this low impedance develops the necessary voltage across the line at point *R*. By adjusting the characteristic impedance of the telescoping section experimentally, reasonably satisfactory coupling has been obtained over approximately a 2-to-1 frequency range with this coupling system.

This coupling method has the advantage of loading the parallel-line tank circuit at a voltage node for any mode of operation of the parallel line. Thus it does not tend to favor certain modes as is sometimes the case with direct coupling (see Sec. 16-11). The unbalance introduced by this type of coupling does not appear to be a serious disadvantage, at least in the case of the split-anode magnetron.

CHAPTER 17

MODULATION OF HIGH-FREQUENCY POWER OSCILLATORS

BY J. G. STEPHENSON AND W. R. RAMBO

I. AMPLITUDE MODULATION

17-1. Requirements for Modulators.—Modulation is discussed, in the following sections, with respect to rather special and limited characteristics and requirements. In general, the modulated devices are self-excited oscillators. The principal objectives are (1) maximum side-band power, regardless of phase shifts, distortion, and incidental frequency modulation, and, in most cases, (2) maximum band width. The first objective permits a high degree of modulation on the positive peaks and plate-current cutoff on the negative peaks, provided that the total power output does not suffer too much. The achievement of band widths of, say, 100 kc to 5 Mc, simply and without sacrifice in total power, involves certain compromises with and limitations by factors such as circuit Q 's and modulating capacitances.

17-2. Fundamental Limitations in Band Width Due to High Tank-circuit Q .—The fundamental limitation encountered in obtaining wide band modulation of oscillators (at video frequencies up to 5 Mc) in the range from 25 to 1500 Mc or so is the excessively high loaded Q of the output tank circuit. By definition

$$Q = 2\pi \frac{\text{energy stored}}{\text{energy lost per cycle}}$$

In the case of a lumped-parameter circuit, shown in Fig. 17-1,

$$Q = 2\pi \frac{\frac{E^2 C}{2}}{\frac{E^2}{2R}} f_0$$

or

$$Q = 2\pi f_0 CR \tag{17-1}$$

where

$$f_0 = \frac{1}{2\pi \sqrt{LC}}$$

Eq. (17-1) may be written

$$Q = R \sqrt{\frac{C}{L}} = \frac{R}{Z_0} \tag{17-2}$$

where $Z_0 = \sqrt{L/C}$, the "characteristic impedance" of the resonant circuit.

The resistance term R depends upon the load resistance, the tank-circuit loss resistance, and a shunting resistance due to the oscillator tube, all in parallel. For optimum operation of a given tube in an oscil-

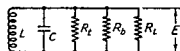


FIG. 17-1.—Lumped-parameter parallel-resonant circuit: L , inductance of tuned circuit; C , capacitance of tuned circuit; f_0 , resonant frequency; E , peak voltage across R ; Δf , frequency band width to half-power points (by definition, $\Delta f = f_0/Q$); R_t , shunting resistance due to oscillator tube, R_b , resistance due to tank-circuit losses; R_L , resistance due to load.

lator circuit, the resistance R is fairly well fixed by the tank impedance required to produce the necessary power output with given plate and bias voltages. In the process of adjusting the oscillator for optimum output, the loading is varied until the resistance R approaches the optimum value. From Eq. (17-1) it is therefore apparent that the shunt capacitance C of a lumped tank circuit must be kept to a minimum to obtain the lowest operating Q and the widest band width. This is one of the

factors favoring the use of variable inductances with only their distributed capacitances for the tank circuits of u-h-f oscillators and amplifiers.

In parallel-line and coaxial tank circuits the same principles apply as to lumped circuits, but must be modified to take into account the distributed nature of the line constants and energy storage. By crude analogy to the lumped-circuit example of Fig. 17-1, it may be correctly inferred that the characteristic impedance Z_0 of a transmission-line tank circuit should be high for wide-band modulation in the fundamental or $\lambda/4$ mode of operation.¹ (This is not necessarily true of higher-mode operation in coaxial lines. For further discussion see Sec. 15-7.)

It is important to appreciate the magnitudes of the loaded tank-circuit Q commonly found in u-h-f oscillators to realize the fundamental limitations in modulation band width which are encountered. Table 17-1 is based on typical examples of various types of oscillators and shows approximate values of operating Q . The calculations are based on a number of estimated quantities, and the Q figures should be taken only as an order of magnitude, but actual modulation performance of the oscillators checks well with the values listed in the table.

The impedance of a high- Q tank circuit falls off rapidly on either side of the resonant frequency. For example, at frequencies $f_0(1 \pm 1/2Q)$ it is 70 per cent of the resonant value, and at frequencies $f_0(1 \pm 1/Q)$ it is only 45 per cent of the value at resonance. This means that the higher the modulating frequency, the smaller the impedance of the tank circuit to side frequencies, and the smaller the percentage of modulation

¹ By fundamental mode is meant the lowest natural frequency of the circuit

TABLE 17-1.—APPROXIMATE OPERATING Q OF VARIOUS ULTRAHIGH-FREQUENCY OSCILLATORS

Oscillator tube	Circuit	Frequency, Mc	Approx. tank impedance, ohms	Approx.* equivalent shunt capacitance, $\mu\mu\text{f}$	Operating Q of tank circuit	Band width to half-power points, Mc
388A push-pull.	Parallel-line tank circuit	300	9,000	4.4	75-80	3.7
368AS push-pull	Parallel-line tank circuit	500	13,200	2.5	105	4.7
8012 push-pull..	Parallel-plate tank circuit	550	7,350	5.1	130	4.0
3C22.....	Coaxial-line	400 ($\lambda/4$ mode)	4,500	10.0	115	3.5
3C22.....	Coaxial-line	900 ($\lambda/4$ mode)	4,500	6.9	175	5.0
3C22.....	Coaxial-line	900 ($3\lambda/4$ mode)	4,500	22.7	580	1.5
2C39.....	Coaxial-line	1200 ($3\lambda/4$ mode)	5,300	19.5	780	1.5

* This figure takes into account the distributed nature of both the capacitance and voltage in resonant-line tank circuits and also includes tube capacitance, strays, and end effects.

of the r-f envelope that can be produced by a given amount of modulating voltage. In practical cases the actual reduction in percentage modulation will depend on the internal impedance of the oscillator or amplifier tube as well as on the impedance of the tank circuit to the side-band frequencies. In most oscillators at frequencies below 1500 Mc, it is extremely difficult to obtain a high percentage of modulation at video frequencies greater than 3 or 4 Mc.

By using an oversized plate modulator, it is possible to increase, to a certain extent, the over-all modulation percentage that can be obtained at modulating frequencies where Q is the chief limiting factor. An alternative method of improving the percentage of modulation with a normal plate modulator is to introduce simultaneous grid modulation of proper phase and magnitude. This method is described in Sec. 17-13.

17-3. Plate-modulating Impedance of Oscillators.—In designing plate modulators it is necessary to know the impedance that an u-h-f oscillator presents to the modulator. Figure 17-2 shows an equivalent circuit of an oscillator as seen by the modulator, with the modulating impedance represented by a resistance R_m in parallel with a capacitance

C_m . The capacitance C_m is determined by the construction of the oscillator and can be readily measured. It may consist of a d-c blocking

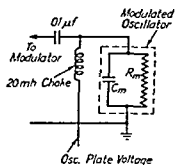


FIG. 17-2—Equivalent plate-modulating impedance of an ultrahigh-frequency oscillator. R_m , equivalent dynamic resistance of the oscillator at modulating frequency. C_m , plate-ground capacitance of oscillator.

capacitance, an r-f by-pass capacitance, or a stray plate-ground capacitance in the plate circuit of the oscillator and may be as low as $15 \mu\text{f}$, or as high as several hundred micromicrofarads in different oscillators. The resistance R_m is approximately equal to $2E_b/3I_b$, at frequencies between 1 and 5 Mc, where E_b and I_b are, respectively, the d-c plate voltage and current of the oscillator when normally loaded. The factor $\frac{2}{3}$ is an empirical one, based on laboratory measurements of the modulating impedance of u-h-f oscillators at video frequencies.

It is important to keep the modulating capacitance at a low value, for obviously it

is much easier to maintain the desired video voltage across a small capacitance. However, if the operating Q of the oscillator is very high, as is true in many cases, it is almost useless to apply 4- and 5-Mc modulating frequencies to the oscillator because of the limitations of tank-circuit Q . Modulating capacitance is therefore not necessarily the most important factor limiting wide-band modulation.

17-4. Impedance of Modulation Chokes.—Single π -chokes commonly used at video frequencies in Heising modulation circuits (see Fig. 17-7) are usually resonant at a frequency below 500 kc and appear essentially like small capacitances at higher frequencies. Figure 17-3 shows measurements of the impedance of two typical modulation chokes. The irregularities are apparently due to subsidiary resonances at frequencies above the principal choke resonance. Most modulation chokes having nominal inductance values of the order of 20 to 60 mh appear like a 5- to $15 \mu\text{f}$ capacitance at modulation frequencies between 750 kc and 5 Mc.

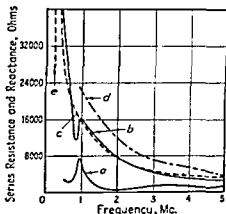


FIG. 17-3—Impedance of two typical modulation chokes at video frequencies: curve a , series resistance of 20-mh choke; curve b , capacitive reactance of 20-mh choke; curve c , reactance of $11\text{-}\mu\text{f}$ capacitor; curve d , capacitive reactance of 60-mh choke (corresponding to $7 \mu\text{f}$); e , principal resonance of 20-mh choke, 272 kc.

17-5. Nonlinearity of Plate Modulation in Ultrahigh-frequency Oscillators.—The usual conditions of operation of plate-modulated u-h-f oscillators are not conducive to modulation linearity for the following reasons:

1. Self-excited u-h-f power oscillators, when fully loaded, will not even oscillate at d-c plate voltages below a certain minimum value (commonly 20 to 40 per cent of the normal plate voltage). This effect is especially pronounced at the higher frequencies where transit-time effects become important.

2. The limitations on percentage modulation imposed by tank-circuit Q cause a departure from linearity at high modulating frequencies.

3. High positive modulation peaks may cause some tendency toward emission saturation with certain oscillator tubes, such as the "doorknob" types. Under rather severe operating conditions, these oscillators will often show a decrease in average r-f power when plate-modulated, as compared with a normal increase in output where no appreciable saturation takes place.

4. A certain amount of incidental grid modulation is frequently present in many u-h-f plate-modulated oscillators, whether intended or not. Even a few micromicrofarads of plate-to-grid capacitance will cause some modulating voltage to appear at the grid when the impedance between the oscillator grid and ground is appreciable at video frequencies. The phase relations between grid and plate-modulating voltages are usually such as to encourage nonlinearity of modulation (see Sec. 17-13).

17-6. Voltage and Power Relations in Modulated Ultrahigh-frequency Oscillators.—Figure 17-4 shows static plate-modulation characteristics of a typical u-h-f oscillator at 300 Mc under various conditions of grid-bias adjustment. Curve a was taken with automatic-resistance grid bias which adjusted itself to be approximately proportional to the plate voltage. This is the static equivalent of the dynamic case in normal plate-modulated Class C oscillators, in which a small amount of variation of the grid leak bias occurs over the modulation cycle, tending to maintain linearity. This bias variation is 180 deg out of phase with the plate-voltage modulation, *i.e.*, the grid potential becomes more negative on the positive modulation peaks.

Curve b was taken with the oscillator bias held constant as the d-c plate voltage was varied. A dynamic counterpart of this case occurs if

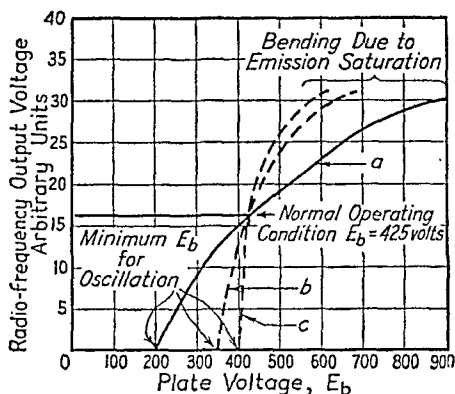


FIG. 17-4.—Static modulation characteristics of push-pull parallel-line oscillator at 300 Mc (WE 388A Tubes).

the time constant of the resistance bias and by-pass capacitance of an oscillator is so long that the bias remains essentially constant over the modulation cycle.

Curve *c* was taken with the bias decreasing as the plate voltage increased. This is the static equivalent of the dynamic case where grid modulation is present in phase with the plate modulation, *i.e.*, the grid potential becomes more positive on the positive plate-modulation peaks.

Consider a plate-modulating voltage superimposed on the normal d-c plate voltage of 425 volts in Fig. 17-4. Even under the conditions of curve *a*, which is the closest approach to linearity, it is apparent that the discontinuity at $E_b = 200$ volts, where oscillations cease, and the tendency toward emission saturation at $E_b = 800$ volts would cause some departure from linear modulation. Under the conditions of curves *b* and *c*, which are commonly encountered in u-h-f oscillators, these tendencies toward nonlinearity are much more pronounced. It can be readily seen that a negative modulating-voltage swing of less than one-third of the d-c plate voltage can cause 100 per cent "down" modulation or even overmodulation at frequencies at which limitations due to tank *Q* are not serious. Where *Q* is the prime limiting factor, however, higher and higher voltages must be applied as the modulating frequency increases in order to obtain a given percentage of modulation until, finally, at sufficiently high modulating frequency, no amount of modulating voltage anywhere in the oscillator circuit will produce a high percentage of modulation.

Under the foregoing conditions, actual modulator output power has no simple relation to the unmodulated carrier power as is the case with linear sinusoidal plate modulation. When maximum total side-band power is required, it is often necessary to use the highest possible peak modulation voltage over the desired band with a given size of modulator and tubes, in order to counteract the discriminating action of high tank-circuit *Q* at side-band frequencies. This may require positive modulation peak voltages greater than the d-c plate voltage of the oscillator.

17-7. Incidental Frequency Modulation of Amplitude-modulated Oscillators.—Plate or grid modulation of self-excited u-h-f oscillators is invariably accompanied by incidental frequency modulation in some degree. The extent of frequency shift in any particular case depends on the slope of the phase-vs.-frequency curve in the feedback network and, hence, on the operating *Q*'s of the tuned circuits, degree of coupling, etc. The resulting frequency excursion in most practical oscillators operating in the 300- to 1000-Mc region seldom extends much beyond 1 or 2 Mc, but may become greater at frequencies above 1000 Mc.

Simultaneous amplitude and frequency modulation is made evident by unequal amplitude of corresponding side bands at frequencies close

to the carrier. Figure 17-5, for example, shows the voltage amplitude of side bands resulting from sinusoidal modulation of a 600-Mc oscillator. Note the difference in amplitude of corresponding side frequencies up to a modulating frequency of about 2 Mc. Even a small degree of incidental frequency modulation results in some reduction in carrier amplitude when modulation is applied.

When noise modulation is applied to u-h-f oscillators, frequency modulation shows up in the spectrum as a "spreading out" of the carrier into a triangular shape instead of the usual carrier "spike." Figure 17-6 illustrates this type of spectrum.

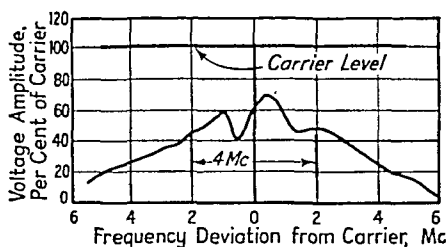


FIG. 17-5.—Amplitude of side bands resulting from sinusoidal plate modulation of a 600-Mc self-excited oscillator, showing effects of incidental frequency modulation.

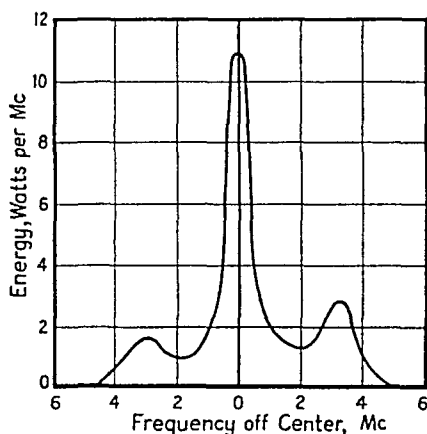


FIG. 17-6.—Typical energy spectrum of noise-modulated oscillator at 300 Mc. Spreading of carrier and unequal amplitude of side-band peaks due to incidental frequency modulation.

17-8. Direct (Impedance) Coupling.—Direct coupling is suitable for coupling a modulator to the plate of an u-h-f oscillator when wide-band video frequencies are involved. Peaking or compensation for the high video frequencies may be carried out as for video voltage amplifiers. An advantage of direct coupling is its simplicity, and a disadvantage is that impedance matching is not applicable.

Modulator tubes are usually pentodes or beam tetrodes and are chosen for suitable power and voltage ratings and for the usual figure of merit g_m/C_i , where g_m is the transconductance and C_i is the output plus input capacitances. A number of modulator tubes may be used in parallel to increase the output obtainable, when the tube output capacitances are smaller than the shunt capacitance across the load. Figure 17-7 illustrates a typical Heising or direct-coupled modulation arrangement for wide-band modulation of a u-h-f oscillator. The use of two modulation chokes permits different plate voltages to be applied to the oscillator and

modulator. Note the use of four modulator tubes in parallel, with small damping resistors in all grid and plate leads to reduce any tendency toward parasitic oscillations.

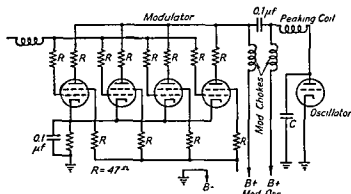


FIG. 17-7.—Typical Heising coupling arrangement for plate modulation of an ultrahigh-frequency oscillator. C , plate-ground capacitance of oscillator

If flatness of modulator response is not of primary importance, very simple methods of frequency compensation are often sufficient. When the oscillator capacitance seen by the modulator is less than about

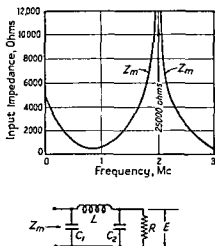


FIG. 17-8—Impedance presented to modulator as function of frequency: R , 5,000 ohms; L , 165 μ h; C_1 , 50 μ f; C_2 , 175 μ f.

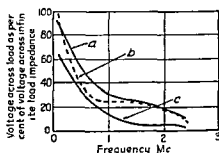


FIG. 17-9.—Voltage across load vs frequency for circuit of Fig 17-8. Modulator-driving voltage somewhat peaked at 2 Mc: curve a , calculated; curve b , measured; curve c calculated for circuit with L shorted out.

50 μ f, it is relatively easy to maintain good modulator response out to 3 Mc or more. When this capacitance is of the order of 200 μ f, compensation in the modulator stage becomes more important than anywhere in the driver stages. By the connection of a simple series peaking coil

between the modulator and load, it is possible to obtain appreciable improvement in the voltage obtainable across the load up to about 2 Mc. The network and input impedance seen by the modulator are shown in Fig. 17-8 for a typical case. The impedance presented to the modulator varies widely because of both series- and parallel-resonance phenomena in the output network. However, as Fig. 17-9 shows, the load voltage drops off gradually with increasing frequency. In the region where a very low impedance is presented to the modulator, the load voltage is sustained by a series-resonance rise in voltage, which causes a high circulating current in the modulator output circuit. The improvement due to the addition of the peaking coil is evident by comparing curves *a* and *c* of Fig. 17-9.

Thus, in oscillators where r-f considerations demand a large plate-ground capacitance, it may still be possible to obtain sufficient plate-modulating voltage for some purposes, provided that the desired upper frequency is not too high and uniform response and phase shift are not required.

17-9. Video Transformer Coupling and Class B Modulators.—In low-power equipments it is generally satisfactory to employ single-ended Class AB Heising modulators, with modulator-tube capability obtained by parallel connection of several tubes sufficient to produce the needed amount of video power. Such modulators operate at relatively low efficiency, and the impedance match between generator and load is seldom good. Because of the relatively low level of operation, this procedure may readily be justified from a design standpoint. However, in equipments requiring considerable amounts of video power, the over-all efficiency of the modulation system becomes important, and the Heising circuit is discarded in favor of transformer-coupled Class B modulators.

It may be shown that the efficiency of an ideal Class B modulator with complex wave forms applied may be expressed as follows:

$$\eta = \frac{E_{pm}}{E_b} \left(\frac{e_{rms}}{e_{av}} \right) \left(\frac{e_{rms}}{e_{peak}} \right) \quad (17-3)$$

where η is the efficiency, E_{pm} the peak plate-voltage swing, E_b the direct plate voltage, and e_{rms} , e_{peak} , and e_{av} are the characteristic values of the applied voltage wave.

In sine-wave operation, the expression for efficiency becomes the familiar value $(E_{pm}/E_b) \times (\pi/4)$. For a complex wave form, periodic or not, values may be assigned to e_{av} , e_{rms} , and e_{peak} . These values for random noise, limited to an arbitrary peak value 10 per cent of the time, show that efficiencies of 50 to 60 per cent may readily be obtained with Class B operation. The design of suitable Class B interstage and output transformers is, however, a problem.

17-10. Basic Design Considerations of Transformers.—The highest frequency that a wide-band transformer will pass is principally determined by the leakage inductance and the distributed capacitance of the windings; the lowest practicable frequency is determined by the magnitude of the primary inductive reactance in relation to the impedance of the power source. In order that a transformer cover as wide a range of frequencies as possible, iron cores are used, which raise the inductance of the primary winding and thereby lower the low-frequency limit for a given set of windings, but generally have little effect upon the high-frequency limit. The higher the permeability of the core, the lower the low-frequency limit of the transformer. Another design consideration

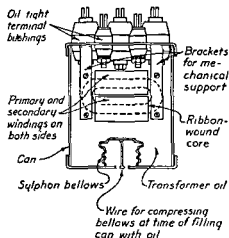


FIG. 17-10.—Oil-immersed wide-band video-frequency transformer.

concerns the variation of core loss with frequency. This variation may be stated in broad terms as follows: a given transformer, at constant applied primary voltage, will exhibit a reduction in core loss as frequency is increased. This fact is important and is true because, for a given number of induced volts per turn, the flux density in the core is an inverse function of frequency. At very high frequencies, the flux density in the core drops to a low value. If a given transformer operates at 12,000 gauss at 60 cycles, for example, the

flux density at constant voltage drops to 12 gauss at 60 kc and 1.2 gauss at 600 kc. It might be expected that at these low flux densities the permeability would be low, and this is experimentally confirmed, even for maximum flux densities as high as 400 gauss at 100 kc.

By employing oil-cooling single-layer coils operating at 150 circular mils per amp, and providing free oil circulation in large containers, it has been found practicable to design video transformers handling up to 400 watts of video noise power in the frequency range 0.1 to 4 Mc, coupling the output of two 250-watt power pentodes to a 50-ohm load. Figure 17-10 shows the construction of a typical oil-immersed video-frequency transformer. In this 40-to-1 frequency range the response is reasonably uniform, and the over-all conversion efficiency of the Class B stage and output transformer is approximately 55 per cent. The complete design procedure is rather involved and cannot be readily presented here. However, a few items can be mentioned. A high ratio of iron to copper (not dielectric) loss must be used, and the smaller the rated transformer power,

the higher the allowable ratio. Designs for frequencies of 100 kc and above may employ flux densities corresponding to 2 volts per turn per square centimeter cross section of core, if the laminations are of the order of 2 mils in thickness. Core materials found satisfactory from a loss standpoint include Hypersil, Monimax, Permalloy ribbon, and B9W4A. To minimize distributed capacitance, coils should be single-layer-wound. On cores with small windows, the inductance will be increased by the core by a factor of approximately 40 at a frequency of 100 kc.

It is believed that video transformers will be used increasingly in the future in a variety of applications. Increased band width in very low-power units awaits the production of core material having high permeability at relatively low flux densities, and in high-power units, core material having high permeability, thin laminations, and low hysteresis loss.

17-11. Grid Modulation of Ultrahigh-frequency Oscillators.—A Class C u-h-f oscillator adjusts itself until the tube just supplies all the losses of the load, excitation, and circuit. Normally this equilibrium is attained with less total r-f plate-voltage swing in relation to the d-c plate voltage than in low-frequency oscillators, as the result of poorer tube efficiency in

converting d-c to r-f power. When grid modulation is applied to an oscillator under these conditions, the applied bias decreases during part of the modulation cycle, *i.e.*, the grid potential becomes more positive than its normal or average value. This tends to increase the amplitude of oscillation somewhat. But inherent limiting factors such as increased load and grid losses and lower plate efficiency prevent the amplitude from building up very far, and under no circumstances would the peak r-f voltage swing exceed the d-c plate voltage. Thus the "upward" modulation is somewhat clipped.

During the downward swing of the grid-modulating voltage, the bias voltage temporarily increases, *i.e.*, the grid potential becomes more negative than its normal or average value, and oscillations tend to die out altogether with an applied signal of very few volts. Thus 100 per cent "down" modulation, or overmodulation, can easily be obtained at frequencies where tank-circuit Q is not an important factor (see Fig. 17-11).

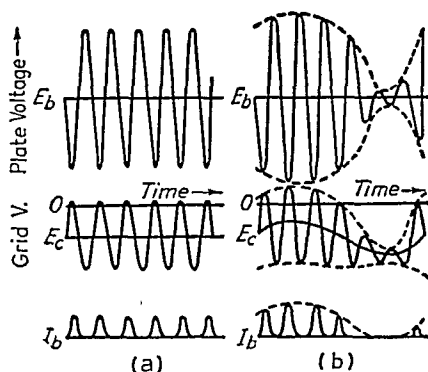


FIG. 17-11.—Approximate voltage relations in an ultrahigh-frequency oscillator with grid modulation: (a) steady-state conditions (no modulation); (b) grid modulation (approximately 20 per cent positive peak modulation, 85 per cent negative peak modulation shown); E_b , d-c plate voltage; E_c , average grid bias; I_b , plate-current pulses; O, cathode potential (reference).

The following characteristics of grid-modulated u-h-f oscillators have been noted:

1. As a result of the action described above, only a small percentage of "upward" modulation can be obtained, *i.e.*, the modulation is principally "downward" and, consequently, distorted.
2. The r-f power output of the oscillator drops as a result of the predominantly downward modulation. This reduction in practical cases is of the order of 30 or 40 per cent at 100 per cent downward modulation.
3. A very small modulator may be used for grid modulation of an oscillator.
4. If the peak grid-modulating voltage greatly exceeds the d-c grid voltage, either pulsing or complete paralysis of the oscillator usually results. Pulsing generally occurs at low modulating frequencies, the oscillator being operative only on the positive modulation peaks. Complete stopping of oscillations at higher modulation frequencies is due to excess d-c bias developed in the grid circuit of the oscillator.
5. The modulation band width appears to be limited by tank-circuit Q exactly as in plate modulation.
6. Grid modulation of u-h-f oscillators often causes a greater tendency toward quenching difficulties, though the troubles are not necessarily prohibitive.

17-12. Cathode Modulation.—Cathode modulation of oscillators, where the modulating voltage is impressed across an impedance in the

cathode circuit with the grid and plate effectively grounded at video frequencies, is essentially grid modulation with a small percentage of in phase plate modulation. This is illustrated in Method I, Fig. 17-12c. Under these conditions the effects of grid modulation predominate, and the behavior is essentially the same as that described for grid modulation.

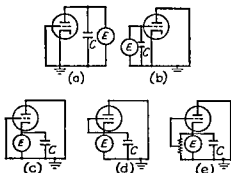


FIG. 17-12.—Elementary equivalent circuits for plate, grid, and cathode modulation of an oscillator or amplifier: (a) plate modulation; (b) grid modulation; (c) cathode modulation, method I; (d) cathode modulation, method II; (e) cathode modulation, method III; E , modulating voltage; C , capacitance across modulating voltage.

this method has the disadvantage of requiring low-capacitance filament transformers and special efforts to minimize the cathode-ground capacitance in the oscillator circuit.

If the grid can be returned (at video frequencies) to a point somewhere between cathode and ground, as shown in Method III, Fig. 17-12e, the

percentage of plate and grid modulation can be varied to obtain any desired combination of plate and grid modulation (see also Sec. 17-13).

17-13. Simultaneous Plate and Grid Modulation of Ultrahigh-frequency Oscillators.—When plate modulation is applied to a u-h-f oscillator, circuit conditions may be such that a small part of the modulating voltage is also applied to the oscillator-tube grid. What would ordinarily be simple plate modulation is then plate-plus-grid modulation. The magnitude of the modulating voltage appearing between grid and

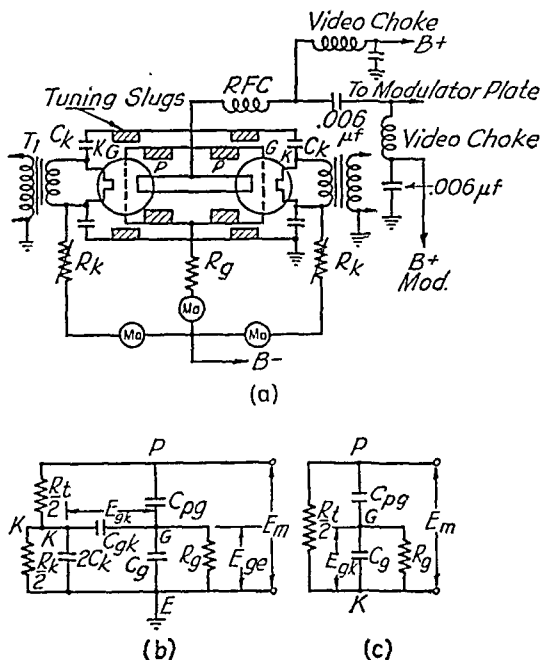


FIG. 17-13.—Push-pull coaxial-line oscillator: (a) actual circuit; (b) exact and (c) approximate equivalent video-frequency impedance network for plate modulation; R_g , grid leak, 500 ohms; R_t , dynamic impedance presented to modulator by one tube (about 10,000 ohms); R_k , cathode resistance, 300 ohms; C_k , cathode by-pass capacitance, .01 μf ; C_{gk} , grid-cathode capacitance, 8 μf ; C_{pg} , total plate-grid capacitance, 32 μf ; C_g , total grid-ground capacitance, 77 μf ; $E_m \approx E_{pk}$, impressed modulating voltage; E_{gk} , modulating voltage appearing between grid and cathode.

cathode and its phase relation to the plate-cathode-modulating voltage in a particular u-h-f oscillator can be predicted with reasonable accuracy by a simple analysis of the equivalent video-frequency impedance network. The conditions that favor the appearance of a large modulating voltage at the grid are a plate-to-grid capacitance of more than about 10 μf , combined with a grid-to-cathode impedance at video frequencies that is appreciable compared with the plate-grid impedance. Other impedances (such as those between cathode and ground) may also influence the results somewhat. Note that capacitive reactances that are very

high at audio frequencies may be quite low at frequencies between 1 and 5 Mc.

A push-pull coaxial oscillator and its equivalent video-frequency impedance network are illustrated in Fig. 17-13. The grid-modulating voltage appears as a result of

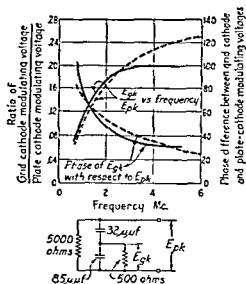


FIG. 17-14.—Magnitude and phase of grid-cathode modulating voltage for push-pull oscillator of Fig. 17-13. Solid lines, measured values; dashed lines, calculated values. Simplified equivalent video-frequency network shown at bottom.

incidental grid modulation is also present. (Magnitude and phase of the incidental grid modulation for curve *a* is shown in Fig. 17-14.) Curve *b* shows the marked decrease in band width that results when the grids of the oscillator are by-passed to ground with a 0.01- μ f capacitance in order to remove all grid modulation.

Figure 17-16 shows the effects of intentionally varying the phase relation at constant magnitude and of varying the magnitude at constant phase of the grid-modulating voltage with respect to the plate-modulating voltage for the oscillator of Fig. 17-13. Note that the phase relation is not critical as long as the grid-cathode-modulating voltage is not more than about 90 deg out of phase with the plate-cathode-modulating voltage. In

impedance network are illustrated in Fig. 17-13. The grid-modulating voltage appears as a result of voltage-divider action of the plate-grid and grid-cathode impedances at modulating frequencies. The magnitude and phase of the grid-cathode modulating voltage calculated from the simplified equivalent circuit are compared with the measured values in Fig. 17-14 and show agreement within the accuracy of measurement of voltages and circuit parameters. (The effect of grid current was neglected in the calculations.) The effects of incidental grid modulation upon the oscillator of Fig. 17-13 are shown in Fig. 17-15. Curve *a* shows the amplitude of the side bands resulting from sinusoidal plate modulation where

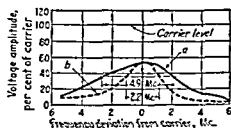


FIG. 17-15—Effect of incidental grid modulation on amplitude of side bands resulting from sinusoidal plate modulation of the oscillator of Fig. 17-13; curve *a*, plate plus incidental grid modulation; curve *b*, plate modulation only; frequency 600 Mc; plate-modulating voltage held constant

this particular case the optimum magnitude of the grid-cathode voltage is about 20 per cent of the plate-cathode voltage. In general, this optimum value will depend on the particular oscillator circuit, tubes, magnitude of the plate-modulating voltage, and d-c voltages. The optimum peak value will usually be found to be of the same order as the d-c grid bias.

The grid-modulating voltage will be found to be appreciable in some oscillator circuits and negligible in others, depending on the physical construction and on the magnitude of various by-pass condensers and other impedances in the oscillator circuit. In some cases it will be found that the video-frequency impedances can be varied appreciably without

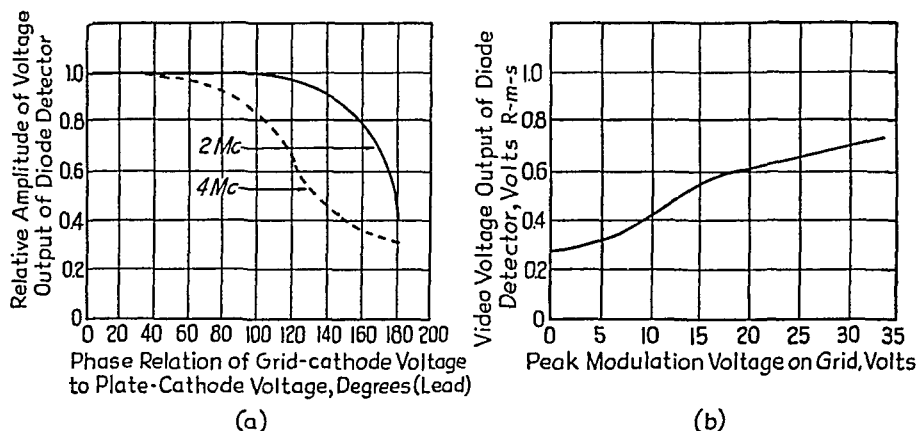


FIG. 17-16.—Effect of varying the magnitude and phase of the grid-cathode-modulating voltage with respect to the plate-cathode-modulating voltage for the oscillator of Fig. 17-13: (a) variation of phase, plate-modulating voltage, 193-volt peak, grid-modulating voltage, 29-volt peak, both held constant; (b) variation of amplitude, phase difference held constant at 0 deg, plate-modulating, voltage, 193-volt peak.

upsetting the r-f characteristics of the oscillator, and thus some measure of control of the incidental grid modulation may be possible. It may also be possible to cause the magnitude of the grid-modulating voltage to increase with modulating frequency, and thus to provide the most grid modulation at the higher video frequencies where limitations due to high loaded-tank-circuit Q are serious and where it is consequently most needed. In other oscillators, as, for example, the single-ended coaxial-line oscillator described in Sec. 15-15 and Fig. 15-15, no grid modulation is present because the grid is grounded and there is a large by-pass capacitance between cathode and ground. It is not possible in this type of oscillator to obtain properly phased grid modulation merely by adjusting circuit impedances. However, simultaneous in phase grid and plate modulation have been obtained by using a separate cathode modulator, tapped from a low-level stage of the main plate modulator.

With the type of cathode modulation shown in Method III, Fig.

17-12e, it is theoretically possible to obtain simultaneous inphase grid and plate modulation and to vary the relative amounts of each type, as described in Sec. 17-12. It is not ordinarily convenient to apply this method to u-h-f oscillators because of the requirement for low cathode-ground capacitance. Also, in grid-separation coaxial oscillators it is usually not convenient to provide a high video-frequency impedance between grid and ground.

With a given size of plate modulator, the percentage of modulation of an oscillator having a high- Q tank circuit ordinarily falls off rapidly as the modulating frequency increases, as is evident from Fig. 17-15. The addition of inphase grid modulation has the effect of increasing the power of the plate modulator and enables a higher percentage modulation to be obtained for a given plate-modulating voltage. This characteristic is especially useful at high modulating frequencies where limitations due to tank-circuit Q are serious. By causing the grid-modulating voltage to increase at the higher frequencies, this method offers the possibility of obtaining more equalized response over a band of modulating frequencies than would be possible when using an ordinary plate modulator with a flat output-vs.-frequency characteristic. In other words, it is a method of compensation that may be more convenient than using a much larger plate modulator.

The improvement in band width due to the application of simultaneous grid modulation is limited. The maximum increase in band width obtained experimentally thus far is about 100 per cent. Modulation obtained in this manner may be nonlinear, particularly at low modulating frequencies where caution must be used to avoid overmodulation. With proper adjustment of the relative magnitude and phase of the grid modulation, however, it has been possible to obtain an appreciable increase in the percentage of modulation at frequencies between 1 and 5 Mc without serious distortion of the modulating voltage. (Envelope with sinusoidal modulation was observed on wide-band high-speed-sweep oscilloscope.)

17-14. Qualitative Explanation of the Effects of Simultaneous Grid and Plate Modulation.—Analysis of the static modulation characteristics of u-h-f oscillators (as described in Sec. 17-6 and Fig. 17-4) shows that when the grid-bias potential varies 180 deg out of phase with the plate voltage the modulation tends to be most nearly linear. If the bias is held essentially constant throughout the modulation cycle, or if it varies in phase with the plate modulation, there is a tendency for the oscillator to cut off rapidly on the "down" modulation peaks and to approach a higher level on the positive peaks, *i.e.*, it requires less plate-voltage swing to obtain a given degree of modulation. Available evidence indicates that this tendency still holds in the dynamic case where high Q is the main factor limiting wide-band modulation.

Figure 17-17c is the same as Fig. 17-17b except that inphase grid modulation is applied. The grid is driven positive at the same time that the plate goes positive in the modulation cycle. This causes larger pulses of plate current to flow and tends to produce a greater voltage across the tank-circuit impedance on the positive modulation peaks than in Fig. 17-17b. On the negative modulation peaks, the grid is driven negative at the same time that the plate voltage is reduced. Plate-current pulses become very small and may finally cease altogether, allowing the r-f oscillations in the tank circuit to die out at the maximum possible rate. Hence inphase grid modulation, by increasing the amplitude of the side-frequency components of the plate current, increases the side-frequency voltages in the tank circuit somewhat, despite the reduction in impedance at the side frequencies.

Certain factors inherently limit the extent of this improvement in modulation as the modulating frequency becomes very high. The tank-circuit impedance to side bands falls off rapidly at high modulating frequencies, cathode emission may limit the increase in plate-current pulses on the positive modulation peaks, a larger angle of flow of grid and plate-current pulses causes increased losses on positive modulation peaks, and the finite rate of decay of r-f oscillations becomes serious at high modulating frequencies. At modulating frequencies above approximately $2.5f_c/Q$, these limiting tendencies prevent obtaining a high percentage of modulation, though a noticeable improvement is still possible through simultaneous grid modulation.

Though no study of the problem has been made, there is no apparent reason why the foregoing principles should not apply to Class C amplifiers as well as to self-excited oscillators.

II. FREQUENCY MODULATION OF HIGH-FREQUENCY POWER OSCILLATORS

Practical cases arise in which it is desirable to frequency-modulate power oscillators operating at high frequencies and thereby avoid the use of amplifying and frequency-multiplying circuits. The design and operation of circuits for direct modulation are discussed in this section. The same end result often can be achieved by the more familiar methods of multiplying, in frequency and power, the output of a modulated low-level low-frequency oscillator or by amplifying the modulated output of a high-frequency velocity-modulation tube. The final choice of methods must depend on the particular requirements to be satisfied as regards fundamental frequency, power, modulated band width, linearity, and allowable incidental amplitude modulation.

17-15. Common-grid Reactance-tube Circuits.—The common-grid reactance-tube circuit is shown schematically in Fig. 17-18a. The r-f

driving voltage is developed across an impedance between cathode and grid in a circuit in which the grid terminal (which may operate at r-f ground potential as in Fig. 17-18a) is common to both driving circuit and output circuit.

The configuration of the common-grid reactance-tube circuit lends itself to incorporation, physically, in v-h-f and coaxial u-h-f oscillators. Electrically, it permits the use in simple circuits of the available high-frequency triode tubes and has characteristics that tend to reduce or eliminate the incidental amplitude modulation that often accompanies the frequency-modulation process. Thus, while many conventional reactance-tube circuits can be applied to some extent to u-h-f use, most of this section will be devoted to a discussion of the properties of the common-grid circuit.

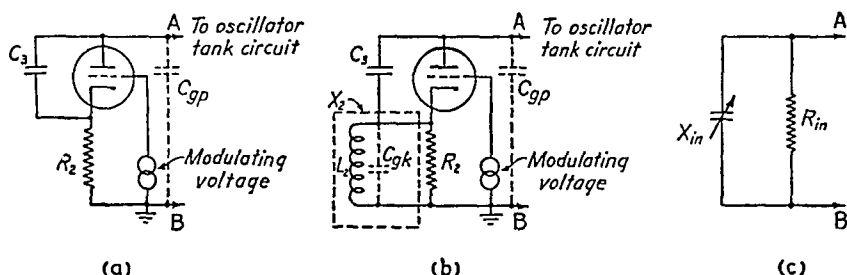


FIG. 17-18.—Triode common-grid reactance-tube circuit: (a) Case I, C_{pk} small so that in the parallel combination, X_{cok} can be neglected in comparison with R_2 ; (b) Case II, C_{pk} paralleled by an inductance, L_2 , such that X_2 is an inductive reactance equal in magnitude to X_{cok} ; (c) the variable reactance seen looking into the terminals A - B.

A *reactance tube* is a tube in which the plate current flows out of phase with the applied r-f plate voltage. This flow of plate current is usually controlled in amplitude and phase by a combination of an r-f grid voltage, as developed in a phase-splitting network, and the desired audio or video modulating voltage. The reactance simulated by the tube (the ratio of the applied r-f plate voltage to the reactive component of plate current) is a function of the modulating voltage and when introduced into an oscillator tank circuit affects the oscillator frequency and so causes frequency modulation that is a function of the reactance-tube modulating voltage.

There is a variable dissipation of oscillator power in a simple reactance tube attributable to two r-f plate-current components. One is the current in the tube plate-cathode impedance as a result of the r-f voltage applied to the plate. The second is the grid-controlled inphase plate-current component resulting from the fact that the grid voltage is generally not exactly 90 deg out of phase with the plate voltage. This loss of oscillator power is variable with modulation and results in what can be called *parasitic amplitude modulation*.

It is possible to minimize or completely eliminate parasitic amplitude modulation by a number of circuit modifications, as described in several published papers.¹ All are similar in that the shift in grid-voltage phase is made slightly greater than ± 90 deg with respect to the applied plate voltage so that, in effect, the reactance tube is capable of supplying energy to the system in the proper amount to compensate for its own variable internal losses. That is, one plate-current component, in phase with the applied plate voltage, is balanced by the real component of the grid-voltage-controlled plate current 180 deg out of phase with the applied voltage. The reactive component of the grid-voltage-controlled plate current produces the desired reactance effect.

The basic common-grid circuit is shown in Fig. 17-18. In this discussion, the following circuit and tube parameters will be used:

X_3 , the capacitive reactance appearing between the plate and cathode terminals of the reactance tube. This may be the reactance of the interelectrode capacitance alone or augmented by electrostatic feedback. $X_3 = 1/\omega C_3$, where C_3 is the equivalent capacitance between plate and cathode.

X_2 , the reactance appearing between the cathode and grid terminals. It may be the reactance of the interelectrode capacitance alone or that reactance paralleled by the reactance of a tuning circuit.

R_2 , the resistance appearing between cathode and grid terminals and paralleling X_2 .

X_m , the variable reactance (see Fig. 17-18c) seen looking into the terminals $A-B$ which, when bridged across an oscillator tank circuit, produces frequency modulation. The fixed interelectrode capacitance of the tube from plate to grid is assumed lumped in the oscillator tank circuit.

R_m , the equivalent resistance (see Fig. 17-18c) paralleling the terminals $A-B$ and X_m . This resistance produces a loading of the oscillator tank circuit; the higher the resistance, the less the loading. To reduce incidental amplitude modulation, it is necessary to hold this value constant or at least to minimize any variations.

μ , amplification factor of the tube

r_p , plate resistance of the tube, ohms

g_m , transconductance of the tube, mhos

The so-called *phase-splitting network* is composed of X_3 in series with the parallel combination of X_2 and R_2 . The driving voltage is developed across R_2 .

¹ For example, see SHAEFFER, C. F., Frequency Modulator, *Proc. I.R.E.*, **28**, 66 (1940); TRAVIS, CHARLES, Automatic Frequency Control, *Proc. I.R.E.*, **23**, 1125 (1935); BUTLER, F., Reactance-valve Frequency Modulator, *Wireless Eng.*, **20**, 539 (1943).

17-16. Analysis of the Common-grid Circuit.—An analysis of the common-grid circuit can be made under the assumption of Class A operation and the equivalent-plate-circuit theorem. The general equations for the resistance and reactance paralleling an oscillator tank circuit connected across the points AB in Fig. 17-18 are

$$R_{in} = \frac{R_2^2 \frac{(X_3 - X_2)^2}{X_2^2} + X_3^2(1 + R_2 g_m)^2}{\frac{R_2^2 X_3}{X_2^2 r_p} (X_3 + \mu X_e) + R_2 + \frac{X_3^2}{r_p} (1 + R_2 g_m)} \quad (17-4)$$

$$X_{in} = \frac{R_2^2 \frac{(X_3 - X_2)^2}{X_2^2} + X_3^2(1 + R_2 g_m)^2}{R_2^2 \frac{(X_3 - X_2)}{X_2^2} - \frac{R_2 X_3}{r_p} \left(1 + \frac{R_2 X_3 g_m}{X_e}\right) + X_3(1 + R_2 g_m)} \quad (17-5)$$

where X_3 is positive for a capacitive reactance, X_2 is positive for an inductive reactance, and it is assumed that $\mu \gg 1$ and $R_2 \ll r_p$. In the modulation process, the tube transconductance is varied in accordance with the modulating voltage.

Most practical applications of the circuit involve two special cases. Case I (Fig. 17-18a) holds principally at low frequencies where the reactance of the grid-cathode interelectrode capacitance is large compared with the resistance in the phase-splitting network so that X_2 is much larger than R_2 . The frequency to which this condition holds is a function of tube characteristics and may be as high as 20 Mc for small power triodes. The parallel resistance looking into the terminals AB is then

$$R_{in} \approx \frac{R_2^2 + X_3^2(1 + R_2 g_m)^2}{R_2 + \frac{g_m}{\mu} X_3^2(1 + R_2 g_m)} \quad (17-6)$$

When the tube is nonconducting, this reduces to

$$R_{in} \approx R_2 + \frac{X_3^2}{R_2} \quad (17-7)$$

The parallel reactance simulated by the circuit when the tube is conducting is approximately

$$X_{in} \approx X_3(1 + R_2 g_m) \quad (17-8)$$

so that the oscillator tank circuit is paralleled with a capacitance

$$C \approx \frac{C_3}{1 + R_2 g_m} \quad (17-8a)$$

If the quantity k satisfies the relation

$$k = \mu R_2^2 C_3^2 \omega^2 \quad (17-9)$$

Eq. (17-6) can be written

$$R_{in} \approx R_2 \frac{k + \mu(1 + R_2 g_m)^2}{k + R_2 g_m(1 + R_2 g_m)} \quad (17-10)$$

and will behave as follows:

if $k \leq 0.5$, R_{in} will decrease with an increase in g_m from the value at $g_m = 0$.

if $k \geq 1.0$, R_{in} will increase with an increase in g_m , never returning to the value at plate-current cutoff.

if $0.5 < k < 1.0$, R_{in} will have a maximum value, as g_m goes from zero to infinity that is never more than four-thirds the resistance when the tube is nonconducting. Furthermore, at some value of g_m , g_m' , the resistance will pass through the same value it has at $g_m = 0$. (Such a curve is shown in Fig. 17-20.)

The last case is interesting, for if k is chosen according to the relation

$$k = \frac{(1 + R_2 g_m')}{(2 + R_2 g_m')} \quad (17-11)$$

the resistance and oscillator-circuit loading will be equal at tube transconductance values of $g_m = 0$ (the value existing at one end of the modulated band)

and of $g_m = g_m'$ (which may be selected from the tube characteristics to be a value at the other end of the modulated band). Between these end points, the maximum resistance value will never

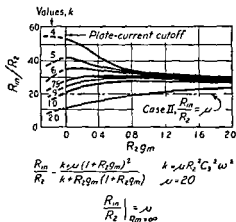


FIG. 17-19.—Variations in oscillator loading caused by modulation of a typical triode common-grid reactance tube.

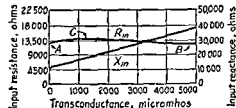


FIG. 17-20.—Characteristics of a triode common-grid reactance-tube circuit calculated from Eqs. (17-8) and (17-10), $k = 0.75$, $\mu = 20$; $f_0 = 13$ Mc; $C_{kp} = 1 \mu\text{f}$; $R_1 = 450$ ohms.

exceed four-thirds of the value at $g_m = 0$ and, depending on the value of k , generally will be much less. A plot of Eq. (17-10) for various values of k in a typical case is given in Fig. 17-19. The ratio R_{in}/R_2 is plotted as a function of the product $R_2 g_m$.

Figure 17-20 illustrates the characteristics of a small triode common-grid reactance-tube circuit calculated for the case where $k = 0.75$. The

resistive component is the same at transconductance values of 0 and 4700 μ mhos. The maximum value of the curve occurs at a value of 1300 μ mhos (point *C* in Fig. 17-20) and exceeds the value at *A* and *B* by about 11 per cent. Variable loading is thus minimized without additional circuit elements.

The solution of Eq. (17-11) requires the assumption of a value for R_2 .¹ This will specify X_3 through Eq. (17-9). With the phase-splitting network constants R_2 and X_3 chosen, an idea of the maximum available reactance change and oscillator loading can be gotten from Eqs. (17-8a) and (17-7). If either is unsatisfactory, a new calculation must be made to obtain revised network constants. The amount of oscillator loading and the maximum effective reactance change are not independent, and a larger frequency-changing effect is obtained at the expense of an increased (though still nearly constant) absorption of oscillator power.

Case II arises if the frequency is such that the reactance of the grid-cathode interelectrode capacitance is not large compared with the phase-splitting resistance R_2 (a condition for Case I) or if it is desired further to lessen variable loading of the oscillator. The cathode-grid circuit then must be paralleled with an inductance of such magnitude that the susceptance component of the total cathode-grid admittance is of opposite sign, but equal in magnitude, to the susceptance of the phase-splitting capacitor C_3 . If the circuit constants are adjusted at the center frequency so that

$$\frac{1}{\omega C_3} = R_2 \sqrt{\mu} \quad \text{and} \quad \omega L_2 = \frac{1}{\omega C_3} \quad (17-12)$$

where ωL_2 is the reactance of the parallel grid-cathode capacitance and external inductance, then

$$X_{in} \approx X_3(1 + R_2 g_m) \quad (17-13)$$

and

$$R_{in} \approx \mu R_2 \quad (17-14)$$

To the extent to which μ remains constant and that Eq. (17-12) can be satisfied, constant oscillator loading is produced by the common-grid reactance-tube circuit shown in Fig. 17-18b. This characteristic would appear in Fig. 17-19 as the dotted curve of constant value. This ideal characteristic can be compared with an actual experimental curve in Fig. 17-21, which shows the output of a small oscillator, frequency modulated over a ± 4 per cent band at about 28 Mc by a triode common-grid reactance tube.

An estimate of the linearity of frequency change as a function of modulating voltage can be obtained from the curve in Fig. 17-22, which is

¹ A typical assumed value for R_2 would be one such that $R_2 g_m' = 1$.

a plot of measured results on a 60-watt experimental transmitter, the circuit diagram of which appears in Fig. 17-23.

Figure 17-24 is a photograph showing the small physical size possible in this unit designed to operate at 50 Mc with a frequency swing of

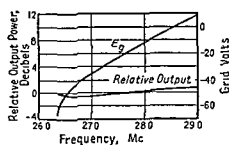


FIG. 17-21.—Typical performance of a small oscillator frequency-modulated by a type-6C4 triode reactance tube adjusted for constant oscillator loading (Case II). Oscillator output at 0 db, 2 watts.

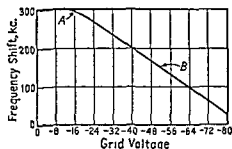


FIG. 17-22.—Measured frequency shift vs. grid voltage of push-pull type-6C4 common-grid reactance tubes modulating a type-829B oscillator (Fig. 17-23). Oscillator output, 60 watts, reactance-tube plate voltage, 200 volts; oscillator plate voltage, 700 volts; operating frequency f_0 , 48 Mc.

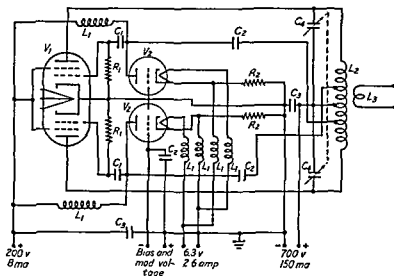


FIG. 17-23.—Frequency-modulated oscillator: V_1 , 829B, V_2 , 6C4; R_1 , 8200-ohm 1-watt carbon, R_2 , 330-ohm 2-watt carbon, C_1 , 10 μf , C_2 , 1000 μf , C_3 , 10,000 μf (ceramicons); C_4 , 2-sec. 15- μf variable capacitor; L_1 , ohmite type-Z1 r-f choke; L_2 , 1 μh , 8T coil l 2 in., d $1\frac{1}{4}$ in.; L_3 , 1T, d $1\frac{1}{4}$ in. Note: phase-splitting capacitance (not shown) formed by C_{pk} .

± 100 kc. Incidental amplitude modulation is negligible, and the required modulating power is less than 1 watt.

The power relations existing in the common-grid circuit are generally similar to those in any conventional reactance-tube circuit, with the

exception that a portion of the power loss shifts from the phase-splitting resistance to the tube itself under modulation even though the over-all circuit loss may remain constant.¹ This is not a particular disadvantage because this additional tube dissipation is a small fraction of the major power loss in normal high-current reactance-tube operation (the dissipation at the reactance-tube plate of the energy supplied by the reactance-tube plate power supply in normal tube operation). The power aspects of reactance-tube operation are discussed in Sec. 17-20. Equation (17-12) can be satisfied at only one frequency in the modulated band, but

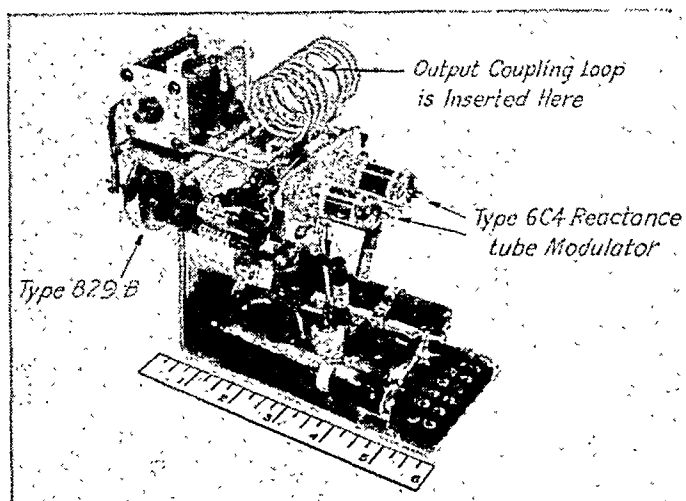


FIG. 17-24.—Photograph of oscillator-modulator test setup (Fig. 17-23).

the circuit frequency sensitivity is not a practical limitation to the operation over wide frequency-modulation band widths.

The low Q of the phase-splitting circuit minimizes any tendency toward parasitic oscillations in the reactance tube on possible resonant frequencies that might exist in this circuit. If X_2 is inductive, as in Case II, no trouble will be experienced with reactance-tube oscillation.

Practical precautionary measures toward the prevention of unwanted oscillations in Case I consist mainly in specifying X_3 several times as large as R_2 and are automatically satisfied when the conditions outlined in Eqs. (17-9) and (17-11) are met.

17-17. The Common-grid Reactance-tube Circuit at Ultrahigh Frequencies.—The common-grid reactance-tube circuit is directly adapt-

¹ Figure 17-20 shows that the reactance-tube circuit input impedance increases with tube transconductance (compare total impedance at points A and B). For a given impressed voltage across terminals A-B in Fig. 17-18a, less current will flow through the circuit, and the loss in the series resistance (I^2R_s) will decrease. The total absorption of oscillator power must remain constant so there must be an increasing amount of dissipation of oscillator power in the reactance tube.

able to coaxial oscillator circuits, but the difficulty of obtaining a precise knowledge of circuit parameters at ultrahigh frequencies makes Eq. (17-12) difficult to apply to quantitative circuit design. In order to achieve peak circuit performance, it is always necessary to make adjustments under operating conditions, usually in the reactance-tube cathode-circuit tuning and in the equivalent plate-cathode capacitance C_3 , which may be adjusted through variations in the amount of external feedback.

It is usually possible to estimate the total capacitance in the tank circuit to be modulated. The desired percentage frequency shift and an

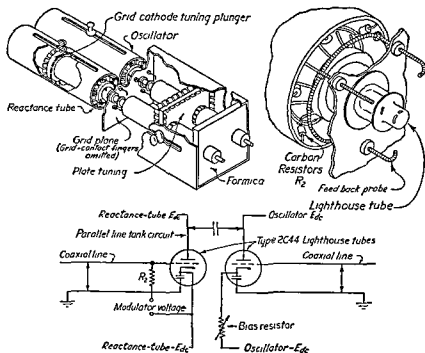


FIG. 17-25 — Ultrahigh-frequency oscillator-reactance-tube unit.

estimate of the maximum value of $R_2 g_m$ in Eq. (17-13) dictates the maximum value of X_3 that will provide the required tank-circuit reactance change. Smaller values of X_3 will provide greater reactance variations, but will also increase the oscillator loading. With $1/\omega C_3$ determined, R_2 can be calculated from Eq. (17-12), and with a knowledge of the oscillator r-f tank voltage, the oscillator power loss can be estimated by evaluating R_{in} in Eq. (17-14).

17-18. A Frequency-modulated Oscillator for Ultrahigh Frequencies. An example of a low-power coaxial-line type of oscillator frequency modulated by a common-grid reactance tube is shown in Fig. 17-25. A 2C44 lighthouse triode is used as a conventional grounded-grid oscillator with plate and cathode circuits tuned by short-circuited sections of trans-

mission lines. The reactance tube is also a type 2C44 triode. The two tubes are placed side by side with the grids in a common plane and the plates tightly coupled together in a common cavity. Blocking capacitors in the plate-short-circuiting plunger allow separate biasing of the grids and, if desired, separate power supplies for the two plate circuits. Internal blocking capacitors for the grid bias are incorporated in the tubes.

The oscillator utilizes conventional capacitive feedback between the plate and cathode cavities. The phase-splitting network reactance X_3 consists of the plate-cathode interelectrode capacitance of the reactance tube augmented by electrostatic feedback (see Fig. 17-18). This feedback is obtained by three wire probes, extending through the grid plane into the plate and cathode cavities and equally spaced on a $1\frac{3}{4}$ -in. circle centered on the tube center line. In the cathode cavity, the probes are linked by a wire loop. In the plate cavity, the probes are bent toward the plate of the tube. The feedback reactance X_3 can be controlled by varying the distance of the probes from the plate. The resistance R_2 , in Fig. 17-18, consists of nine 5600-ohm $\frac{1}{2}$ -watt carbon resistors placed radially about the opening of the reactance-tube cathode cavity. This cavity, a short-circuited section of coaxial transmission line tuning C_{kt} , represents L_2 in Fig. 17-18b. The oscillator cathode line is physically a duplicate of the reactance-tube cathode line, but without the resistance.

Practical realization of the circuit possibilities requires proper tuning of the adjustable elements. The oscillator will tune over more than a 2-to-1 tuning range. The value of R_2 , 622 ohms, was chosen to satisfy Eq. (17-12) for a frequency of approximately 750 Mc, and it is at this frequency that optimum performance occurs. The procedure to satisfy Eq. (17-12) is as follows: The oscillator is tuned to the desired operating frequency with the reactance tube at cutoff. The reactance-tube grid voltage is then adjusted to give the maximum frequency deviation desired; and the cathode line X_2 is tuned to equalize the power output at this point with that obtained at reactance-tube cutoff. This procedure involves shifting the reactance-tube plate-current phase so that the tube delivers more or less real power to the tank circuit—the correct phase being that which enables the tube to neutralize its own losses. Increasing the line length increases the delivery of real power. The amplitude modulation can thus be adjusted to low values over the modulated band width at this particular frequency. Some amplitude modulation will occur when the operating frequency differs considerably from the design frequency. The reactance introduced into the grid-cathode circuit at ultrahigh frequencies by the use of carbon resistors for R_2 is unimportant, as compensating tuning adjustments can generally be made to produce the desired value of X_2 .

Figure 17-26 illustrates the incremental modulation of the type 2C44

oscillator over a frequency band of approximately 0.6 per cent at 734 Mc (± 2.2 Mc) with the output constant within 0.5 db. At the point of maximum frequency deviation, the reactance-tube plate dissipation reached 25 watts. Good linearity of modulation is seen to occur over

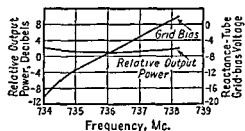


FIG. 17-26.—Static modulation characteristics of an ultrahigh-frequency triode oscillator frequency-modulated by a common-grid reactance-tube circuit. Type-2C44 oscillator: E_{d1} , 500 volts; I_{d1} , 40 ma; I_g , 16 ma. Type-2C44 reactance tube: E_{d2} , 400 volts; I_{d2} , 60 ma; I_g , 11 ma (max). Oscillator power output at 0 db, 6 watts; reactance-tube plate dissipation at 0 volts bias, 25 watts; oscillator resonant frequency at reactance-tube current cutoff, 734 Mc.

ratus with a 25 per cent loss of measured c-w oscillator output in the equivalent parallel circuit resistance R_{L1} . An increased loss of output would accompany wider band modulation.

17-19. Class C Operation of Reactance Tubes.—In modulating power oscillators over wide band widths at high frequencies, it is frequently impossible to set up Class A conditions in the modulator in the practical case, nor is such operation always desirable. It has been found in practice that reactance-tube dissipation for a given modulation can sometimes be halved by operating the reactance tube with a reduced angle of plate-current flow. Another important feature is the fact that the characteristics of the power tubes required to handle the dissipations involved in wide-band modulation of power oscillators are such as to fit Class B or C operation more normally than Class A.

Practically, Class B or C operation requires only an r-f grid-driving voltage of increased amplitude, compared with that required for Class A, and the proper readjustment of grid bias to provide the desired angle of flow of r-f plate current.

17-20. Power Relations in Reactance Tubes.—It is customary to evaluate reactance-tube performance on the basis that the vacuum tube operates in a Class A condition, i.e., that the wave shape of plate current is a reproduction of the wave shape of the signal voltage applied to the grid. This is a valid assumption under most conditions, and circuit performance can be calculated on both a quantitative and qualitative

about 80 per cent of the band. The setup described here operated with a minimum effective loaded Q of 200. It will be shown in Sec. 17-20 that proper tank-circuit design for lower Q operation would reduce the reactance-tube dissipation when modulating a given power over a given percentage band width. The type 2C44 tube plates must be cooled when the oscillator is operating at this power level.

Frequency-modulation band widths in the order of 1 per cent are obtainable with this appa-

basis by viewing the tube as a circuit element whose magnitude can be varied by a modulating voltage. However, when considered on a power basis, the frequency-changing effect of a reactance tube will depend on the percentage of the total oscillator tank-circuit reactive volt-amperes controlled by the tube. There will be a loss of energy at the reactance-tube plate which may be very large in wide-band modulation of power oscillators and, since it must be dissipated at the plate, will be a limiting factor in reactance-tube performance. Class C operation of reactance tubes affords a means of reducing this dissipation without affecting the tube performance as a frequency changer.

In this discussion, the following symbols are used:

E_{\min} = minimum instantaneous r-f voltage appearing across the reactance-tube circuit

E_{dc} = reactance-tube d-c plate-supply voltage

e' = minimum instantaneous plate potential reached during the angle of flow, θ

θ = angle of flow of reactance-tube plate-current pulse

E_m = peak r-f voltage applied to the reactance tube

E_{gm} = maximum instantaneous reactance-tube grid voltage

I_{dc} = reactance-tube d-c plate-current component

I_1 = effective value of reactance-tube fundamental-frequency plate-current component

i_b = instantaneous reactance-tube plate current

The voltage-current phase relations in the Class C reactance tube are indicated in Fig. 17-27. In Class C amplifier operation, tube efficiency is improved because plate current flows at the point of minimum instantaneous plate voltage with resultant low tube loss. In the reactance tube, plate current flows 90 deg away from the E_{\min} value, when the plate voltage is E_{dc} , the d-c supply voltage; thus tube losses are not reduced by the same mechanism as in amplifier operation.

The reactance-tube plate-current pulses can be resolved (if the angle of flow is known) into a d-c component and a fundamental-frequency component in typical Class C manner.¹ The d-c power input to the reactance tube is the product of this d-c component of plate current and the reactance-tube d-c plate voltage. Almost all this power is dissipated in the tube since the plate current, with proper grid-voltage adjustment, normally flows in the wrong phase to deliver or subtract energy from the oscillator tank circuit. The volt-amperes in the reactance-tube circuit are given by the product of the fundamental-frequency plate-current

¹ TERMAN, F. E., "Radio Engineers' Handbook," Sec. 5, Par. 21, p. 444, McGraw-Hill Book Company, Inc., New York, 1943.

component and the r-f tank voltage appearing across the reactance tube and will increase for a given plate dissipation as the angle of flow is

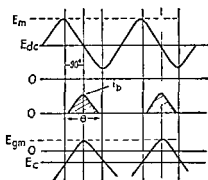


FIG. 17-27.—Approximate phase relations in an ideal reactance tube. Plate voltage and plate current out of phase by 90 deg; plate current and grid voltage nearly in phase. Distortion of plate-current pulse by unsymmetrical plate voltage neglected.

reduced, resulting in a larger frequency shift for a given reactance-tube plate dissipation.

A further improvement in efficiency can be obtained by utilizing the fact that there is no need to maintain E_{dc} equal to or greater than E_{gm} to reduce grid dissipation as in amplifier operation, since these voltages now occur 90 deg out of phase. It is possible, therefore, to drive the instantaneous plate voltage to zero or even to negative values as long as a suitable positive plate voltage exists during the normal angle of plate-current flow.¹ This condition is realized physically simply by lowering E_{dc} to a value that may be even less than the peak r-f plate-voltage swing. The $E_{dc}I_{dc}$ losses are reduced, while the reactive energy controlled by the tube is not affected as long as the r-f plate current and r-f tank voltage remain constant. The amount of reduction in E_{dc} is limited by the need for maintaining a positive plate voltage during the angle of flow, and in most practical cases the reactance-tube plate-supply voltage is set only slightly less than the peak r-f voltage swing in the plate circuit, as shown in Fig. 17-28.

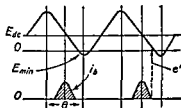


FIG. 17-28.—Reactance-tube operation with reduced plate-supply voltage. Minimum plate voltage, during angle of flow given by the expression: $e' = E_{dc} - E_m \sin \theta/2$

An approximate calculation of circuit performance may be made in a manner similar to Class C amplifier analysis. Some fundamental differences are obvious, and it is helpful to make simplifying assumptions and approximations beyond those normally made in an amplifier analysis. For instance, it is generally assumed that the plate current flows 90 deg out of phase with the tank voltage and that the space-current pulse is

¹ The minimum instantaneous plate potential e' reached during the angle of flow under ideal conditions is

$$e' = E_{dc} - E_m \sin \frac{\theta}{2} \quad (17-16)$$

symmetrical about its peak value.¹ The errors introduced by these assumptions will vary from case to case but do not destroy the value of the analysis as a preliminary aid to setting up a reactance-tube circuit.

The peak reactance-tube plate dissipation is roughly proportional to the Q of the tank circuit of the modulated oscillator, the total output power of the oscillator (useful output plus r-f circuit losses), and the peak percentage frequency deviation.² The dissipation under conditions where the frequency varies above and below an average value (the point B in Fig. 17-22) will, of course, be less than the dissipation at the peak swing A . The Class B or C operation, however, is set up for the expected peak swing away from the oscillator frequency when the reactance tube is nonconducting. It is obviously desirable to maintain the loaded operating Q as low as possible in the modulated circuit in order to reduce dissipation.

An idea of reactance-tube dissipation can be obtained from Table 17-2.

TABLE 17-2.—CLASS C REACTANCE-TUBE PLATE DISSIPATION

Fundamental frequency, Mc	Oscillator tube	Reactance tube	Oscillator output, watts	Average deviation, %	Frequency swing	Average reactance-tube plate dissipation, watts
30	Single section of type 829-B	Single section of type 829-B	10	+5	$\pm 1\frac{1}{2}$ Mc	19
30	Single section of type 829-B	Single section of type 829-B	35	+2	± 600 kc	22
50	Push-pull type 829-B	Push-pull type 6C4	65	+0.25	± 100 kc	10
800	(Water-cooled) 2C44	(Water-cooled) 2C44	7	+0.25	± 2 Mc	12

The loss of power in the reactance-tube network resistance, though constant with proper design, is appreciable when large driving voltages are required. This dissipation increases rapidly as the angle of flow is reduced, since more driving voltage must be developed across the cathode-circuit resistance and the suitability of Class B or C operation to a particular case will depend on the power loss that can be tolerated.

¹ The instantaneous plate potential at the peak of the reactance-tube space-current pulse is E_{dc} , the plate-supply voltage.

² This is so because the total stored energy in the circuit is a function of the product $P_o \times Q$. A change in frequency requires a change in circuit stored energy proportional to the percentage frequency deviation, and this change is brought about by a greater or smaller flow of reactance-tube plate current.

When a reactance-tube frequency-modulated oscillator is to be used at a single fundamental radio frequency, it is possible frequently to couple the radiation resistance of the antenna into the reactance-tube phase-splitting network as a portion or all of the required network resistance and thereby eliminate the normal loss of useful r-f power in that circuit, element. This requires an adjustment of the over-all phase-splitting network impedance to load properly the oscillator tank circuit. However, such an arrangement cannot be used with the common-grid circuit as the grid-cathode driving voltage in this circuit is not constant since the current through the cathode impedance varies with modulation.

The phase-splitting network may be any one of the several types providing for optimum phase adjustment. Circuit-element magnitudes must be such that the appropriate grid driving voltage is derived from the applied oscillator r-f voltage.

The selection of network element values on the basis of a Class A analysis, as is usually done, is a good qualitative start to the final values and may be entirely satisfactory under some circumstances. Thus, though the modulated oscillator circuits discussed in the section on common-grid reactance tubes were set up on a Class A analysis basis, the actual operation was found improved, in some cases, by Class C operation. The curve plotted in Fig. 17-22 was obtained under Class C conditions.

For most applications it is necessary to amplitude-modulate the reactance tube since it must produce frequency modulation in accordance with an amplitude-modulated signal voltage. This is conveniently done in the grid circuit in a manner similar to the grid modulation of Class C amplifiers and with similar modulating-voltage requirements. Here the similarity ends, for there is little or no variation in oscillator or reactance-tube plate-voltage swing under modulation, nor are there the usually severe restrictions on grid-voltage swing imposed by linearity considerations as in amplitude modulation.

17-21. Frequency Modulation by Reactance Switching.—If an actual physical reactance of constant value is periodically inserted into the frequency-determining circuit of an oscillator for a fraction of each oscillator r-f cycle, it will be found that the oscillator frequency is a function both of the value of the reactance and of the portion of the r-f cycle during which it is inserted. By controlling the moment in the cycle and the duration of the insertion in accordance with the modulation, frequency modulation may be attained. A contrast of this with the more conventional methods shows that a physical reactance of constant value is placed in the oscillator tank circuit for a portion of each r-f cycle instead of an electronically simulated reactance for the entire cycle. In one case, the length of time the reactance is inserted is made a function of the

modulating voltage, while in the other the actual magnitude of the reactance is determined by the modulating voltage.

The basic circuit is given in Fig. 17-29. The required reactance X is necessarily an inductance when used with a vacuum-tube switch because of phase relations and unidirectional current flow in a vacuum tube. Either a lumped or a linear circuit element may be used, connected to the oscillator tank circuit through the switch S , which must be an electronic device to permit the moment and duration of switching to be a function of voltages of both modulation and oscillation frequencies.

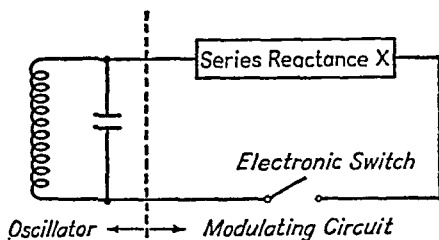


FIG. 17-29.—Basic reactance-switching circuit.

If, for example, the electronic switch closes during that portion of the cycle in which the instantaneous voltage applied to the tube exceeds some arbitrary amplitude, such as E in Fig. 17-31, the time of flow of current during each r-f cycle and, hence, the quantity of reactive current and the volt-amperes controlled are made a function of the modulation voltage. The flow of current through the modulating circuit is essentially 90 deg out of phase with the oscillator tank-circuit voltage when the series react-

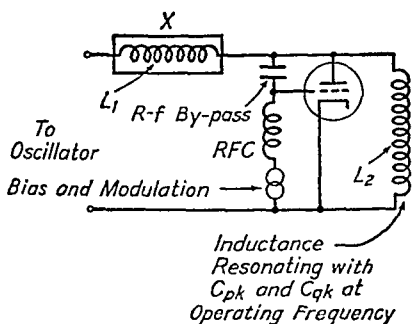


FIG. 17-30.—Practical circuit with high- g_m triode tube. (L_2 may be omitted at frequencies where $X_{Cpk+Cpk} \gg X_{L1}$.)

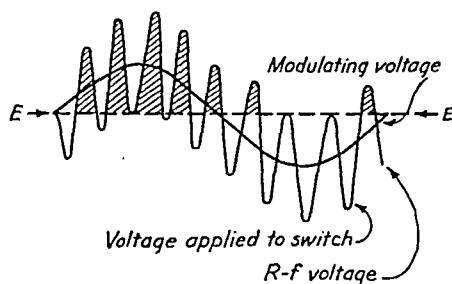


FIG. 17-31.—Variation of r-f current during a modulation cycle. Switch closed during shaded portion of r-f cycle.

ance X in the modulation circuit is large compared with the series resistance presented by the switch during the time of flow. This is true in the practical case at low frequencies, since the minimum resistance of the switching tube is then small compared with the reactance of a suitable circuit element and the branch reactances of the resonant oscillator tank circuit. This is not true, of course, during the time the tube is changing from the "off" to the "on" state, as it must pass through a resistance value comparable to the inserted reactance. This period can be kept a very small

fraction of the r-f cycle so that loading is minimized. Voltages producing excessive grid currents are not applied to the grid element of the triode tube in Fig. 17-30, because as soon as a large current starts to flow, the voltage drop in the series reactance subtracts from the tank-circuit voltage to limit that applied to the tube itself.

An ideal switch would be one that goes from a condition of infinite resistance to one of zero resistance. This is unobtainable in practice because the frequency of operation in most cases dictates the use of a high-vacuum tube as a switching device. Either a diode or a multi-electrode tube connected in a cathode-follower circuit is a usual solution, each having certain advantages with respect to the other, but both having the disadvantage of a finite switching time and an appreciable minimum resistance during the time of current flow in the circuit.

The input admittance of a triode connected as shown in Fig. 17-30 is approximately

$$Y = (g_m + k_p) + j\omega(C_{kg} + C_{kp}) \quad (17-17)$$

where g_m = tube transconductance

$k_p = 1/r_p$

C_{kg} = grid-cathode capacitance

C_{kp} = plate-cathode capacitance

Y = input admittance of the tube looking into the terminals A-B

The conductance is relatively independent of frequency, while the susceptance that shunts this conductance (unimportant at low frequencies) is frequency sensitive.

The impedance of the tube interelectrode capacitances limits the impedance the switch offers at current cutoff and must be tuned out at very high and ultrahigh frequencies.

While the minimum series resistance presented by the vacuum-tube switch during the time of current flow is fairly independent of frequency, the series reactance X must decrease with an increase of frequency to be effective in shunt with the decreasing branch reactances of the normal oscillator tank circuit. The reactance value in ohms is no longer much larger than the switch resistance at ultrahigh frequencies, and the current flow is no longer approximately 90 deg out of phase with the applied voltage. The resultant oscillator loading increases as the frequency is raised and as the ohmic reactance value of X is decreased.

The total loading of an oscillator by a modulating circuit of this type is a complicated function of a number of effects. Practically, the absorption of oscillator power increases with the operating frequency but, at a given fundamental frequency, varies with the incremental frequency

shift, as shown in Fig. 17-32. At low frequencies, tuning ranges as high as 2:1 may be swept without appreciable loss of oscillator power or incidental amplitude modulation. In the v-h-f and u-h-f ranges, react-

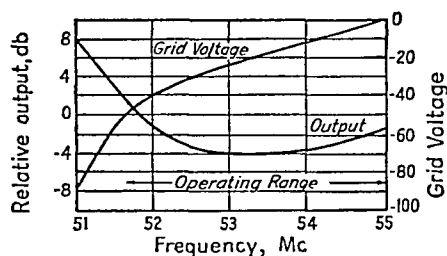


FIG. 17-32.—Characteristics of a 6C4 triode oscillator frequency-modulated by the circuit of Fig. 17-30.

ance-tube circuits compensated for parasitic amplitude modulation are generally preferred to the circuit described in this section because they load the oscillator less.

CHAPTER 18

GENERAL DESIGN CONSIDERATIONS OF HIGH-POWER ULTRAHIGH-FREQUENCY TRIODE AND TETRODE OSCILLATORS

BY W. G. DOW AND GUNNAR HOK

18-1. Introduction.—The design of high-power triode or tetrode oscillators for ultrahigh frequencies differs from the design of low-power oscillators mainly in the respect that larger values in equipment and manpower are at stake in the laboratory, as well as in the field, and that cut-and-try methods are not so easy and attractive as for small oscillators. A more thorough understanding and careful analysis of the factors that influence the efficiency, frequency stability, and modulation properties of u-h-f oscillators are, therefore, of considerable value in dealing with high-power oscillators. For that reason this chapter outlines the modifications necessary in applying the theory of Class C oscillators in the u-h-f field.

18-2. General Equivalent Circuit of an Oscillator.—Fundamentally all sine-wave oscillators can be considered as made up of three parts, a

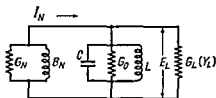


FIG. 18-1.—Equivalent circuit of an oscillator.

nonlinear admittance (G_N, B_N in Fig. 18-1), the real part of which is negative, a resonant system (C, L, G_0 in Fig. 18-1), and a load conductance (G_L), or, if the load is reactive, a load admittance (Y_L). The admittance of a circuit element is said to be non-linear if the fundamental component

of the current through the element is not directly proportional to the sinusoidal voltage component of the same frequency across it, or if the phase difference between these two components is not independent of their amplitude, or both.

In most oscillators in the microwave field the resonant system has distributed constants, *i.e.*, is formed by resonant lines and cavities. For high- Q systems it is, however, possible to represent very accurately each resonance mode within a small frequency range at resonance by an equivalent circuit with lumped constants L, C , and G_0 . Let it be assumed that the frequency of the fundamental component of the current I_N through

the resonant system is the same as the resonance frequency of this system (*i.e.*, $B_N = 0$). In this case the current I_N and the voltage E_L are in phase, and the shunt impedance offered to I_N is simply the reciprocal of the sum of G_0 and G_L . If, for the purpose of this discussion, the Q of the cavity is taken to be the ratio of circulating r-f current to the current through the conductances (as is true in lumped-constant circuits), the loaded Q of the tank circuit is

$$Q_L = \frac{B_c}{G_0 + G_L} = \frac{B_L}{G_0 + G_L} \quad (18-1)$$

where the shunt capacitive and inductive susceptances are equal and can be expressed, respectively,

$$B_c = 2\pi fC \quad (18-2)$$

$$B_L = \frac{1}{2\pi fL} \quad (18-3)$$

Equation (18-1) can be rewritten as

$$Q_L = \frac{1}{G_0 + G_L} \sqrt{B_c B_L} = \frac{1}{G_0 + G_L} \sqrt{\frac{C}{L}} \quad (18-4)$$

Similarly, the unloaded Q is obtained by removing the load conductance G_L

$$Q_0 = \frac{1}{G_0} \sqrt{\frac{C}{L}} \quad (18-5)$$

Equation (18-4) is important because it gives the relation between two important characteristics of the loaded resonator, its shunt impedance $1/(G_0 + G_L)$ at resonance and its band width to modulated waves as represented by the loaded Q . It will therefore be necessary to refer back to this equation when the design considerations for the tank circuits of oscillators are discussed.

18-3. Conditions for Oscillation.—The conditions required for steady-state oscillation are

$$\Sigma B = 0 \quad (18-6)$$

$$\Sigma G = 0 \quad (18-7)$$

These two relations determine the possible frequencies of oscillation and the corresponding amplitudes of oscillation.

In the case of a triode or tetrode oscillator the nonlinear admittance is formed by a feedback network and an amplifier, as shown in Fig. 18-2*a*. A portion γE_L of the voltage across the resonant circuit is applied to the input of the amplifier tube and causes an equivalent electromotive force $E = \mu \gamma E_L$ in the output circuit of the tube, which has an internal imped-

ance z_p . The fraction γ can in general be positive, negative, or complex. The net current delivered by the amplifier tube is

$$I_N = \frac{E_L(1 - \mu\gamma)}{z_p} \quad (18-8)$$

and consequently the nonlinear admittance

$$Y_N = \frac{-I_N}{E_L} = G_N + jB_N = + \frac{1 - \mu\gamma}{z_p} + Y_f \approx + \frac{1 - \mu\gamma}{z_p} \quad (18-9)$$

where Y_f is the input admittance of the feedback network. Y_f is, in general, linear and can be thought of as part of the resonant system.

For tetrode oscillators with high plate impedance the equivalent circuit in Fig. 18-2b is more suitable.

Instead of the voltage generator E with impedance z_p , a current generator I with an admittance y_p is used to represent the internal plate circuit of the tube. The following relations are obtained from Fig. 18-2b:

$$I_N = I - E_L y_p = E_L(\gamma y_m - y_p) \quad (18-10)$$

where y_m is the transadmittance of the amplifier

$$Y_N = \frac{-I_N}{E_L} = -\gamma y_m \left(1 - \frac{y_p}{\gamma y_m}\right) + Y_f \quad (18-11)$$

or

$$Y_N \approx -\gamma y_m \quad (18-12)$$

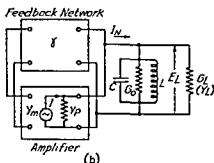
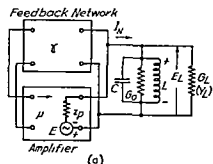


FIG. 18-2—Equivalent circuits of triode and tetrode oscillators.

if y_p and Y_f are negligible in comparison with y_m .

18-4. Frequency Pushing.—In an ideal triode or tetrode oscillator G_N is a function of the amplitude of oscillation but independent of frequency, while B_N may be a function of frequency but independent of the amplitude of oscillation. Such an oscillator is ideal in the sense that the frequency will not change if the amplitude is varied by changing the applied voltages, and the amplitude will not change if the frequency is varied by changing the tuning of the resonator.

At medium frequencies it is possible to realize these conditions μ and z_p are real and independent of frequency, so that B_N will be zero at all

amplitudes and G_N independent of frequency if γ is also real and at least approximately independent of frequency within the frequency range of the oscillator.

At ultrahigh frequencies the conditions are quite different, because of the finite transit time of the electrons. All changes in the flow of electrons in the tube will lag the voltage that causes them. This means that both μ and z_p will be complex. The nonlinear input admittance of vacuum tubes is generally appreciable at these frequencies and must be considered as part of the feedback network. This will tend to make γ nonlinear as well and dependent as to phase and absolute value on transit time effects in the grid circuit. It will be shown later that Y_N must be approximately real in order to obtain good efficiency. It is in general possible to meet this condition also at ultrahigh frequencies by adjustment of γ , but variations in amplitude of oscillation will upset this balance. In other words, at ultrahigh frequencies the condition $B_N = 0$ does not remain satisfied at all amplitudes of oscillation. The fact that variations in amplitude will cause variations in the susceptance B_N is responsible for the so-called *frequency pushing (electronic tuning)*, i.e., the variation of the frequency of oscillation with input current (direct plate current) in u-h-f oscillators.

Frequency pushing is a serious drawback in u-h-f oscillators. Even at medium frequencies, where this phenomenon is not nearly so prominent, the use of high-power modulated oscillators has been abandoned because of their poor frequency stability. Instead, it is the well-known standard practice to use small crystal-controlled constant-amplitude oscillators followed by a number of power-amplifier stages, one of which is modulated.

A similar course of development can be expected in the microwave field, but at the present time power amplifiers are not available above 1000 Mc for any appreciable power. Therefore the problem of frequency pushing, which exists for magnetrons and klystrons as well as for feedback oscillators, is still of considerable importance.

One consequence of frequency pushing is that the average or carrier frequency generated by the oscillator will depend on the applied voltages; it will therefore be subject to variations unless these voltages are carefully stabilized. Furthermore, amplitude modulation will always be accompanied by frequency modulation. This means that the modulated transmission will occupy an unnecessarily broad frequency channel. The congestion in the frequency spectrum is not yet a serious problem in the microwave range, but the wider frequency band transmitted reduces the signal-to-noise ratio at the receiver and thus requires for the same signal-to-noise ratio a larger transmitter power than would purely amplitude-modulated transmission.

An example of the magnitude of the frequency pushing of a high-power oscillator can be given from modulation tests with "resnatron" tetrode oscillators, to be described in the Chap. 19. Frequency variations of the order of 1 to 3 Mc were obtained at a carrier frequency of 500 Mc when the plate current was varied between a low value and full power.

The possibility of using frequency modulation presents a more favorable picture than amplitude modulation. The inverse of pushing, *i.e.*, amplitude variations caused by the frequency variations, will in general be of negligible magnitude.

In order to discuss the factors that determine frequency pushing and the measures that can be taken to reduce pushing to a minimum, it is convenient to consider a tetrode oscillator in the terms of Fig. 18-2b and Eq. (18-12). A tetrode is chosen because of its definite advantages for u-h-f operation, which will be discussed in Sec. 18-9, and also because Chap. 19 will deal with a remarkable development in the field of tetrode oscillators. The arguments presented here can, however, with a slight modification also be applied to triodes.

In the expression for the nonlinear admittance given by Eq. (18-12), only the absolute value, but not the phase, of Y_N will vary with the amplitude of oscillation if y_m is nonlinear but real, *i.e.*, if the transit time of electrons is negligible. Consequently, as was pointed out above, B_N and the frequency of oscillation can be maintained constant if Y_v is real ($B_N = 0$).

When the transit time of the electrons in the tube is not negligible in comparison with one cycle of the oscillation, y_m is not only nonlinear but also complex. Both the absolute value and phase of Y_N then vary with the amplitude of oscillation. The feedback factor can be adjusted so that $B_N = 0$ at one amplitude of oscillation, but when the amplitude varies B_N will not remain zero, because γ is linear, or if the grid losses of the tube are appreciable, nonlinear but according to another—and quite different—relation from y_m . The transit time and the phase angle necessarily vary with the applied voltages; for a given tube very little can be done to reduce these variations. Reduction of the frequency variation *can be obtained if the phase angle of γ varies rapidly with frequency* in such a way as to compensate for the phase change in y_m . This, however, requires among other conditions a high- Q feedback circuit and low grid losses, conditions that are seldom realized in high-power high-efficiency oscillators.

If the variations of B_N are more or less determined by the factors mentioned above, it is still possible to reduce the frequency pushing by designing a resonant system that requires a very small change in frequency to give a susceptance equal to B_N but of opposite sign. The susceptance of the resonant circuit in Fig. 18-2 is

$$B = \omega C \left(1 - \frac{1}{\omega^2 CL} \right) = \omega C \left(1 - \frac{\omega_0^2}{\omega^2} \right) = \omega C \left(1 + \frac{\omega_0}{\omega} \right) \left(1 - \frac{\omega_0}{\omega} \right) \\ \approx 2\delta\omega C \approx 2\delta \sqrt{\frac{C}{L}} \quad (18-13)$$

where $\omega_0 = 1/\sqrt{CL}$ and $\delta = (\omega - \omega_0)/\omega$.

To ensure as little frequency modulation as possible, the tank circuit of the oscillator therefore should be designed with as small $\sqrt{L/C}$ as possible; at the same time the loaded Q should be kept low enough to permit transmission of the amplitude-modulation side bands. According to Eq. (18-4) this leads to a low shunt impedance. This is in conflict with efficiency considerations, as will be shown in a following section. For a given tube it is thus necessary to make a compromise among frequency stability, modulation band width, and efficiency.

18-5. Mode Selection.—When the resonant system has more than one resonance mode, it is rather important to predict and control which mode the oscillator will excite. The general principles of mode selection can be derived by an extension of the fundamental analysis outlined above in Sec. 18-4.

Free oscillations in linear resonant systems have the form

$$E_n \phi = A_n e^{\alpha_n t} \sin(\omega_n t + \phi_n) \quad (18-14)$$

where A_n and ϕ_n are constants dependent on the initial conditions. For a simple circuit consisting of a capacitance C , an inductance L , and a conductance G

$$\alpha_n = \alpha = -\frac{G}{2C} = -\frac{\omega}{2} \cdot \frac{G}{\omega C} \approx -\frac{\omega}{2Q} \quad (18-15)$$

$$\omega_n = \omega = \sqrt{\frac{1}{CL} - \frac{G^2}{4C^2}} = \omega_0 \sqrt{1 - \frac{1}{4Q^2}} \approx \omega_0 \quad (18-16)$$

In a linear system with a number of resonance modes one oscillation of the form of Eq. (18-14) can exist for each mode, and the attenuation and the frequency of each oscillation can be interpreted in terms of the constants of the equivalent circuit of the corresponding mode according to Eqs. (18-15) and (18-16).

The relations given above are not valid for nonlinear systems, but it is a fair approximation to consider all physical systems linear for infinitesimal amplitudes. We can thus discuss the initial instants when the oscillations start building up. Empirical knowledge of the behavior of nonlinear oscillators enables us to state at least qualitatively what happens when the oscillations reach finite amplitudes.

If conditions favorable to oscillation exist in a system with a nonlinear negative conductance and a number of resonance modes, the thermal

agitation pulses will excite all the modes. All modes for which the exponent α_n is positive will start building up exponentially at the rate

$$\begin{aligned}\alpha_n &= \frac{G_{Nn} - G_n}{2C_n} = \frac{\omega_n}{2} \sqrt{\frac{L}{C}} (G_{Nn} - G_n) \\ &= \frac{\omega_n}{2} \left(\frac{1}{Q_{Nn}} - \frac{1}{Q_n} \right)\end{aligned}\quad (18-17)$$

where the subscript n as before denotes an arbitrarily defined mode number and N refers to the nonlinear admittance as in Fig. 18-1. The quantity $Q_N = \omega C/G_N$ has been introduced by analogy with $Q = \omega C/G$.

The mode with the largest exponent will very soon predominate. Usually, but not necessarily always, the nonlinearity will have the effect of gradually reducing the negative conductance for the predominant oscillation as the amplitude increases, until equilibrium is established. The "saturation" of the nonlinear element by the predominant oscillation suppresses the oscillations in the other modes. In other words, the exponents of the other oscillations, averaged over a cycle of the predominant oscillation, are negative, and those oscillations will therefore die out. It is a general rule for nonlinear sine-wave oscillators that under steady-state conditions only one frequency is generated at a time. When two or more frequencies, except harmonics, are present in the output of an oscillator it is often due to rapid periodic shift between two modes. It may also be due to a modulation or cyclic interruption of the oscillations. In the latter case the oscillator may be said to oscillate with two different frequencies, but the frequencies are usually of quite different order of magnitude, and the equivalent circuit and mechanism of oscillation are also different, so that it is more correct to regard the device as two different oscillators with some common elements.

The expression for the exponent given by Eq. (18-17) states that three factors determine the initial mode selection: the frequency, the Q of the mode, and Q_N , the "negative Q " due to the nonlinear element. High-frequency high- Q modes are favored if Q_N does not vary unfavorably with frequency. For triode and tetrode oscillators it was shown in Sec. 18-3 that the negative conductance is determined chiefly by the feedback factor γ and the transadmittance y_m (Eq. 18-11). In order to design an oscillator to operate in one particular mode of the resonant system, we should thus make the feedback network support the desired mode and discriminate against all others, particularly those of high Q . If the feedback network does not discriminate sufficiently against one particular mode, it is often possible to change the resonance frequency of this mode in such a way that sufficient discrimination in the feedback network is obtained.

A common example of a system with two or more modes is a resonant-line tank circuit, tunable and driven from a point at a fixed distance from one end of the line, as shown in Fig. 18-3. The driving point *AA* is the point on the line where the electrode system of the tube is situated. In order to eliminate factors unimportant to the argument it is convenient to assume that G_{Nn}' , the value of the negative conductance G_{Nn} in Eq. (18-17) measured at the point *AA*, remains constant when we tune the resonator and that the load conductance G_n measured at a voltage antinode also remains constant. In determining the exponent of rise α_n , all the circuit constants are referred to the voltage antinode, and the following expression is obtained:

$$\alpha_n = \frac{\omega_n}{2} \sqrt{\frac{L}{C}} (a_n G_{Nn}' - G_n) \quad (18-18)$$

where a_n is a transformation constant that is equal to unity when *AA* coincides with the voltage antinode and equal to zero when *AA* coincides with a node. When the tuning of the resonant line is changed from the state indicated in Fig. 18-3*a* through the state shown in Fig. 18-3*b*, the exponent of rise α_1 for the fundamental mode will gradually increase, while the exponent α_2 for the second mode will gradually decrease. If α_2 is largest initially, α_1 will at some point equal α_2 and from then on remain the larger until α_2 rises again to the same value on the other side of the voltage node. Thus as long as the resonant line is tuned to a relatively low frequency, the oscillator may operate in the second mode, but as the tuning is changed toward higher frequencies it will prefer instead to operate in the fundamental mode. This analysis refers only to the mode selection at the start of oscillations and does not directly apply to the tuning of an oscillator during continuous operation. It is obvious, however, that when a_n approaches zero at the voltage node of a certain mode, oscillations can neither build up nor be sustained in that mode.

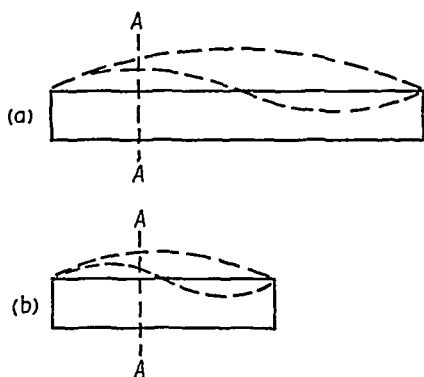


FIG. 18-3.—Excitation of different modes in a resonant line.

18-6. Circuit Efficiency and Electronic Efficiency of Oscillators.—In the preceding sections the conditions under which an oscillator will oscillate and the circumstances that govern mode selection in oscillating systems with more than one resonance mode have been outlined. The next five sections will deal with a third important topic concerning oscillators,

viz., their efficiency. It is customary to consider separately the efficiency of two processes in the oscillator: (1) the conversion of d-c power to a-c power due to the interaction between the electrons and the field in the interelectrode space, which gives the *electronic efficiency* or *electron-interaction efficiency*; and (2) the transfer of the a-c power through the tank circuit to the output terminals of the oscillator, which gives the *circuit efficiency*.

Most of the discussion here will be concerned with the electronic efficiency. An expression can easily be found for the circuit efficiency in terms of the constants of the equivalent circuits in Figs. 18-1 and 18-2. The power delivered to the output terminals will be represented by $E_L^2 G_L$ and the losses in the tank circuit by $E_L^2 G_0$. Consequently, the circuit efficiency η_c is given by

$$\eta_c = \frac{G_L}{G_0 + G_L} \quad (18-19)$$

The electronic efficiency is conventionally taken as the *plate efficiency*, i.e., losses in grid and screen circuits are not included. The grid driving power will consequently have to be accounted for as an additional loss conductance across the tank circuit or as part of the output.

18-7. Electronic Efficiency of Class C Oscillator at Medium Frequencies.—Before u-h-f oscillators are discussed, it may be well to review briefly the factors that determine the plate efficiency of Class C oscillators at medium and low frequencies, where the time of transit of the electrons between the electrodes is negligible compared with the period of oscillation.

In any Class C amplifier or oscillator, or similar electronic device that generates a-c power, whether at low, high, or ultrahigh frequency, good electronic efficiency will be obtained only if electrons accelerated by the d-c field have little or no energy when they return to the metallic circuit by striking an electrode.

When transit time is negligible, three factors will determine whether or not this condition is satisfied in a triode or tetrode oscillator:

1. The plate swing. The instantaneous value of the plate voltage must at its low extreme approach the cathode potential as closely as possible without causing a sharp reduction in plate current. The crest value of the alternating component of the plate voltage should be 80 to 90 per cent of the direct plate voltage.

2. The angle of flow. Plate current should flow only when the instantaneous plate voltage is close to its minimum; the angle of flow should therefore be considerably less than half a cycle.

3. The phase angle between the fundamental components of plate current and plate voltage. The peak of the current pulse should coincide as closely as possible with the minimum of the instantaneous plate voltage.

Terman has pointed out¹ that the shape of the current pulse has within wide limits negligible influence on the efficiency. The output power is, of course, affected by size and shape of the pulses and, consequently, by the relation between grid drive and grid bias, by emission capability of the cathode, and by the removal of electrons from the electron stream by control and screen grids.

For a particular power output there is an optimum value of the shunt impedance offered by the tank circuit. When current and voltage are in phase, *i.e.*, at tank-circuit resonance, the shunt impedance is the reciprocal of the total conductance $G_0 + G_L$, and the optimum value occurs when the fundamental component of the plate current multiplied by the shunt impedance gives the optimum plate swing. If the shunt impedance is lower than optimum, the plate swing will be smaller and the efficiency will be lower because the electrons will hit the plate with more energy. If the shunt impedance is higher than the optimum value, on the other hand, the current leaving the cathode may be reduced by the low or negative plate voltage, and electrons approaching the plate may be "rejected" for the same reason. The electronic efficiency may remain high, since this efficiency does not take into account control-grid and screen-grid losses, but the output power will drop. In a tetrode the overall efficiency will be reduced if the rejected electrons go to the screen.

The requirements, as far as tank-circuit design is concerned, encountered so far are (*cf.* Sec. 18-4) as follows:

1. $\frac{1}{2\pi\sqrt{CL}} = \text{desired frequency of oscillation}$
2. $G_L + G_0 = \text{reciprocal of optimum shunt impedance for best electronic efficiency}$
3. $G_L = 2 \text{ to } 5 \times G_0 \text{ for reasonable circuit efficiency}$
4. $\sqrt{L/C}$ should be as small as possible, for frequency stability
5. $Q_L = \frac{1}{G_0 + G_L} \sqrt{\frac{C}{L}}$, a value determined by required band width for modulation

As G_0 cannot be made smaller than a certain value given by shape and size of the resonant system, there are five conditions to be satisfied by means of only three variables, L , C , and G_L , which is in general impossible. A compromise will have to be made according to the importance of the various requirements for each particular application.

A large band width with reasonable shunt impedance, for instance, requires a relatively large $\sqrt{L/C}$, which necessarily gives poor frequency stability. High shunt impedance and high frequency stability require

¹ Terman, F. E., and W. C. Roake, Calculation and Design of Class C Amplifiers, *Proc. I.R.E.*, 24, 629 (1936).

small G_L and $\sqrt{L/C}$, which will give poor circuit efficiency and narrow band width.

18-8. Electronic Efficiency of the Ultrahigh-frequency Grounded-grid or Grid-separation Oscillator.—One of the outstanding features of modern u-h-f tetrodes and triodes is that their electrodes are designed to form continuous parts or extensions of the resonant lines or cavities that form the oscillator circuit. Any cavity wall of which the grid forms a part must necessarily separate the anode from the cathode. The grounded-grid circuit, where resonant lines or cavities are connected between cathode and grid, and between grid and plate, is therefore the natural circuit for these tubes.

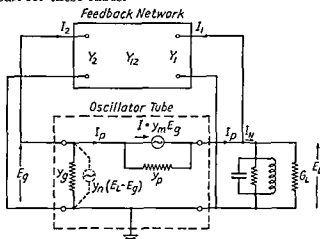


FIG. 18-4—Equivalent network of a grounded-grid oscillator.

An equivalent circuit of the grounded-grid or grid-separation oscillator is shown in Fig. 18-4, drawn in the manner of Fig. 18-2b. The oscillator tube is represented by a grid-cathode admittance y_g , a transadmittance y_m , and a plate-cathode admittance y_p . The feedback network is represented by its input and output self-admittances Y_1 and Y_2 and its transfer admittance Y_{12} . The cathode-grid cavity thus forms part of the admittance Y_2 . From this equivalent circuit the following equations are obtained:

$$I = E_g y_m \quad (18-20)$$

$$I_p = I - (E_L - E_g) y_p \quad (18-21)$$

$$I_1 = I_p - I_N = E_L Y_1 + E_g Y_{12} \quad (18-22)$$

$$I_2 = E_L Y_{12} + E_g Y_2 \quad (18-23)$$

$$I_2 = -(I_p + E_g y_g) \quad (18-24)$$

and the nonlinear admittance

$$Y_N = \frac{-I_N}{E_L} = Y_1 + y_p + \frac{(y_m + y_p - Y_{12})(Y_{12} - y_p)}{y_m + y_p + y_g + Y_2} \quad (18-25)$$

The large admittances in this expression are y_m and Y_2 . Consequently, as a rough approximation

$$Y_N \approx Y_{12} \frac{1 - \frac{Y_{12}}{y_m}}{1 + \frac{Y_2}{y_m}} \approx Y_{12} \frac{1}{1 + \frac{Y_2}{y_m}} \quad (18-26)$$

Allowance for the finite transit time of the electrons has been made by using complex admittances y_o , y_m , y_p for the tube constants. Since the oscillator is operating with grid current, a fourth tube constant exists, the transadmittance y_n representing the influence of the plate voltage on the grid current. The corresponding current generator has been shown with broken lines in Fig. 18-4. In tetrodes, however, this transadmittance is small and will be neglected here.

The feedback circuit shown in Fig. 18-5 gives an example of a practical arrangement for an u-h-f oscillator. Here

$$\begin{aligned} Y_{12} &= -j\omega C_f \\ Y_2 &= Y_k + j\omega C_f \\ Y_1 &= j\omega C_f \end{aligned}$$

where C_f is the feedback capacitance and Y_k is the admittance of the cathode-control-grid cavity. The properties of this circuit at medium frequencies can be judged from Eq. (18-26). y_m is then real, and Y_2 will have to be a capacitive susceptance. Because of the real term in the denominator, Y_N cannot be made a pure negative conductance. The phase condition for good electronic efficiency can thus not be satisfied, as the current I_N (or I_p) is never exactly in phase with E_L . Therefore this circuit cannot be expected to give good efficiency at medium frequencies.

The operation of this circuit at ultrahigh frequencies will be discussed in the following sections, after a description of the motion of the electrons in triodes and tetrodes at frequencies where transit time becomes important.

18-9. Electron Flow in Triodes and Tetrodes at Ultrahigh Frequencies.—The term *transit time* refers to the time required for the flight of an electron through the spaces between the successive electrodes. It is usually described by stating the *transit angle*, i.e., the fraction of a cyclic period in degrees or radians during which the entire flight or some specified portion of it takes place. Thus one may talk of the cathode-to-grid

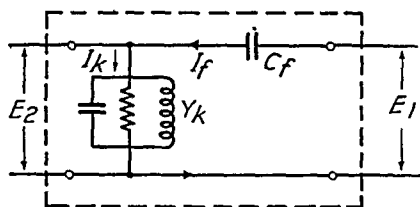


FIG. 18-5.—Feedback network of a grounded-grid oscillator.

transit angle, which measures the time in radians required for an electron to travel from the cathode surface to the plane of the grid. However, electrons that leave the cathode at one moment of the cycle have transit angles different from those departing at another moment. It is usually possible to choose some representative or average value of a suitable parameter and describe the transit angle in terms of that value. For example, the peak value of the grid-plane potential relative to the cathode or the average value of this potential, taken over the angle in which its

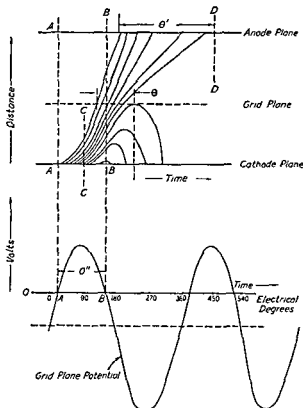


FIG. 18-6.—Distance-time diagram for the flight of electrons in a triode.

instantaneous value is positive, may be used to calculate a characteristic cathode-to-grid transit angle.

In Fig. 18-6 the flight of the electrons in a triode is shown in a qualitative fashion. The ordinate of the upper diagram represents distance from the cathode; the position of the grid and anode are indicated by horizontal lines. The abscissa represents time in degrees of a cycle. The slope of a curve representing the motion of an electron will consequently indicate the velocity of the electron. The sine-wave curve in the lower diagram describes the variation with time of the grid-plane potential of the tube relative to the cathode. At the moment AA the

grid-plane potential passes upward through zero, so that at that moment electrons will begin to leave the cathode. At the moment *BB* the grid-plane potential passes downward through zero, so that from then on no electrons leave the cathode until the moment *AA* is repeated again in the next cycle.

Between the moments *AA* and *BB* all electrons in flight are accelerated toward the grid and pass the grid plane at their maximum velocity. After the moment *BB* the electrons will experience a retarding field during part of their flight and pass the grid plane at progressively lower velocity or return to the cathode. All electrons leaving the cathode before the moment *CC* pass the grid plane; the others are turned back. The first-mentioned group crosses the grid plane during the time corresponding to the angle indicated as θ . The distance-time curves are shown as straight lines between grid and anode. This would be correct only if there were no potential difference, which is obviously not so, but this oversimplification does not substantially alter the conclusion to be drawn.

It might seem proper to consider θ the angle of flow of the plate current. Such is not the case, however, for as long as electrons are traversing the space from grid to anode, current is flowing through the external circuit between these two electrodes, and in the diagram electrons are arriving at the plate as late as at the moment *D*. It may therefore be more nearly correct to call the interval θ' on the diagram the angle of flow. The value of θ' certainly is affected by the assumption that no electric field exists between grid and anode, but the described effect of spreading or *debunching* of the electrons, caused by their different velocities at their entrance into this space, is a real one. A comparison between the shapes of plate-current pulses with and without this effect appears in Fig. 18-7. As the plate current tails off asymptotically to zero in the lower picture, it is a little difficult to define when the angle of flow ends.

This diagram is a convenient basis for a discussion of the various effects produced on the operation of the oscillator by the finite transit time of the electrons.

The most direct consequence is the lag of the electron stream with respect to the alternating grid voltage. Study of Fig. 18-6 shows that the current pulse between cathode and grid lags this voltage by a little

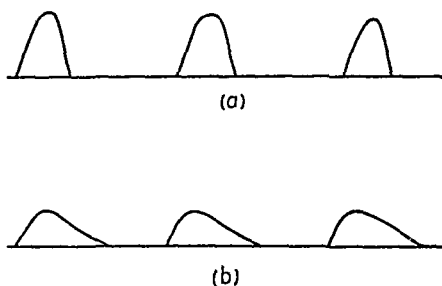


FIG. 18-7.—(a) Typical plate-current pulses when transit time is unimportant; (b) form of plate-current pulse to be expected when debunching effect resulting from unequal grid-traverse velocities is pronounced.

less than half the average cathode-grid transit angle, and the lag of the current between grid and plate is larger by half the average grid-anode transit angle. The transit angle in the cathode-grid space is larger because of the low velocities at the cathode; the effects of the corresponding time lag are consequently observed at lower frequencies than any other transit-time effect.

The angle of flow is affected in two different ways: (1) by the *debunching* of the stream of electrons between grid and plate, caused by the gradual decrease during each pulse of the velocity with which the electrons traverse the grid plane, and (2) by the relatively long *interaction time*, i.e., the time of flight between grid and plate. Both effects tend to increase the angle of flow. To understand the effect of the prolonged interaction time it is necessary only to remember that the plate current associated with any electron begins to flow when that electron passes through the plane of the previous electrode and continues to flow until the electron reaches the anode. It is here assumed that the previous electrode, grid or screen, acts as a perfect electrostatic screen; if this is not true, some plate current will start to flow as soon as the electron leaves the cathode.

In tetrodes the conditions are somewhat more favorable. Between the grid and the screen all electrons will be subject to a very strong accelerating field. When they pass the screen they will have much higher velocity than at the grid plane in a triode, and the relative difference between their velocities will be negligible. In the screen-plate space they will meet a retarding field, so that they will reach the plate with as low velocity as in a triode.

The transit angle between grid and screen may be readily taken into account in this analysis. The r-f impedance between grid and screen should be low so that no appreciable r-f voltage exists between them. The grid-screen transit angle then appears only as part of the phase difference between the ground-to-cathode current ($-I_2$ in Fig. 18-4) and the plate-to-ground current (I_p in Fig. 18-4) and contributes to the susceptance components of the tube constants y_p , y_m , and y_o .

The improvements gained by the introduction of the screen are (1) reduced phase difference between ground-to-cathode and screen-to-ground current, (2) reduction of the angle of flow by elimination of debunching, (3) reduction of the angle of flow by reduction of the interaction time in the screen-plate space.

18-10. Effect of Transit Time on Efficiency.—The qualitative description of the electron motion in triodes and tetrodes given above furnishes some background for a discussion of the efficiency of these tubes as oscillators under conditions of appreciable transit time.

The conditions for good efficiency are not so easily stated as in the

lower frequency oscillators. The instantaneous voltage on the plate is no longer a measure of the kinetic energy of an electron at the impact, because the electron has been moving through a changing field. The kinetic energy of the electron is proportional to the time integral of the electric field in the direction of its path, not to the line integral of the electric field at the instant of impact. This circumstance changes the requirements as to plate swing and phase angle between the a-c plate voltage and the current pulse, but the importance of a small angle of flow remains unchanged.

If it is assumed that all electrons pass the screen or grid plane with the same velocity, a plot of the kinetic energy of the electrons at the impact expressed in electron volts will look qualitatively as shown in Fig. 18-8. The difference in phase position and voltage amplitude between this curve and the instantaneous plate voltage increases with increased interaction time. There is still an optimum plate-voltage swing, but its value is appreciably larger than in a medium-frequency oscillator. The instantaneous plate voltage may very well become negative for a brief interval without causing electron rejection at the anode. The optimum phase angle between the plate-current pulse and the a-c plate voltage is not exactly 180 deg but probably does not differ very much from this value. The reason for the difference is that the electrons, in a tetrode at least, spend more time in the retarding field close to the anode, where their velocity is lower. The transit angle of the electrons that suffer the largest retardation is therefore not centered at the instant of minimum plate voltage but displaced somewhat to the left of this instant in the diagram.

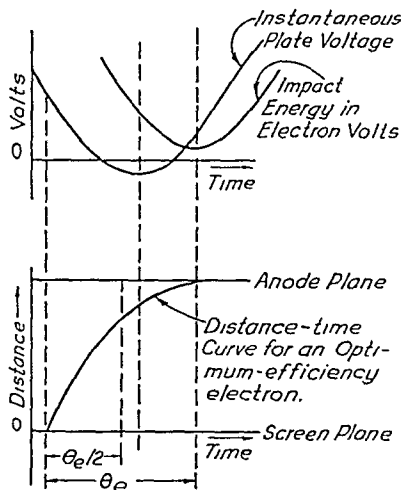


FIG. 18-8.—Electron transit and impact energy.

This discussion indicates that the plate-voltage swing and the phase difference between plate voltage and plate current are for best efficiency subject to restrictions of the same type as at lower frequencies. The optimum plate swing may be 10 to 30 per cent higher owing to the transit-time effect, and the phase angle will differ somewhat from 180 deg, but there seems to be strong evidence, experimental as well as theoretical, that an optimum exists in both these respects. In circuit design, margin for adjustment by experiment should therefore be provided so that approximately optimum conditions can be established in operation.

The considerations regarding tank-circuit design presented in Sec. 18-7 in relation to shunt impedance, band width, circuit efficiency, and frequency stability apply also to u-h-f oscillators, with the quantitative modifications mentioned in the preceding paragraph.

The phase relations in a grounded-grid oscillator under conditions of appreciable transit time can be studied with aid of the analysis in Sec. 18-8 and illustrated by vector diagrams. Equation (18-26) gives an approximate idea of how the phase angle between plate voltage and plate current depends on transit time, as the nonlinear admittance Y_N has the

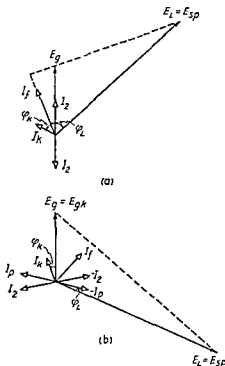


FIG. 18-9.—Vector diagrams for grounded-grid oscillators: (a) negligible transit time; (b) large transit time.

same phase angle. Y_{12} is in general a negative susceptance, and the phase difference between I_p and $-I_2$ is, at least in tetrodes, relatively small. Y_N will then be approximately a pure negative conductance, if the phase difference between Y_2 and y_m is somewhat larger than $+90$ deg, so that the real part of Y_2/y_m will just cancel the number one in the denominator. When this is possible, the denominator can be made very small, and a large negative conductance can be obtained even with a small feedback transadmittance Y_{12} . When the transit-time lag of the plate current is small, a large feedback transadmittance is required, and it may also be difficult to approach the optimum phase condition.

Figure 18-9 shows vector diagrams of a grounded-grid oscillator with the feedback circuit of Fig. 18-5, (1) when transit time is negligible, and (2) when a transit-time lag of 90 to 120 deg exists. For clarity, different arrowheads have been used for voltage vectors and current vectors. The notations refer to the schematic diagram in Fig. 18-4. The ground is used as reference point, which explains the fact that the current I_p changes phase by 180 deg or a little more in passing through the grid. The feedback network in Fig. 18-5 has been assumed in determining the feedback current I_f and cathode-cavity current I_k . ϕ_L and ϕ_k are the phase differences between voltage and current in the anode cavity and cathode cavity, respectively. A study of the top diagram reveals clearly what we have already proved algebraically, *i.e.*, that ϕ_L cannot in this circuit be zero when transit time is negligible. The bottom diagram demonstrates the much better phase relations obtained with appreciable transit time. A much smaller feedback current than the one shown in the diagram would be sufficient to keep ϕ_k below 90 deg.

18-11. Back-heating.—In the description of the motion of the electrons between cathode and grid in an u-h-f triode or tetrode oscillator, in connection with Fig. 18-6 in Sec. 18-9, it was pointed out that a considerable portion of the electrons that leave the cathode fail to pass through the grid and are turned back to the cathode. Some of these electrons will acquire an appreciable velocity and hit the cathode with considerable energy. This is undesirable for two reasons: (1) it increases the conductance component of the grid-cathode admittance and consequently increases the losses in the r-f system, (2) it increases the temperature of the cathode. The latter effect is generally referred to as *back-heating*. If the transit time cannot be reduced, the loss has to be accepted as inevitable, and the cathode has to be designed so that it can take the punishment. High-temperature cathodes with adequate heat transfer are required for tubes of high power. As long as the back-heating power is smaller than the normal heating power, the temperature can be adjusted, manually or automatically, by reducing the heating power until the filament or heater resistance is back to normal.

18-12. Loading Due to Secondary Emission.—A similar phenomenon sometimes takes place between the anode and the screen or grid in low-voltage tubes. Secondary electrons released from the anode may be accelerated by the r-f field before they hit the anode or the screen and may thus load the plate cavity. The effect may be particularly puzzling if the electrons go to the anode, because then the d-c readings of plate current and screen current do not reveal that secondary emission is taking place. The symptoms are increased heating of the plate and low efficiency. The fact that the shunt impedance is reduced further affects the efficiency. As this chapter deals primarily with high-power tubes, a

brief mention of this effect is enough. In high-voltage high-power tubes secondary emission does not present a problem. Measurements that bear this statement out will be presented in the next chapter.

18-13. Trend in Operating Conditions at Progressively Higher Frequencies.—In conclusion it may be of interest to discuss the general trend of the operating conditions as the triode and tetrode oscillators are adapted for progressively higher frequencies. It is evident from the preceding sections that the transit time is a limiting factor; the efficiency must drop fatally if the average transit angle approaches 180 deg because the angle of flow becomes excessive and the back-heating extremely severe. There are two ways to maintain a reasonable transit angle as the frequency is increased: (1) by reducing the spacing of the electrodes, (2) by increasing the operating voltage. The first-mentioned possibility is strictly limited by spark-over difficulties and manufacturing limitations. Increased voltage on the other hand leads to increased current and power, because it is in general not possible to increase the shunt impedance of the tank circuit very much without ruining the circuit efficiency and the band width. From this point of view, therefore, it should be easier to build high-power than low-power oscillators at very high frequencies. It should be pointed out, however, that the difficulties associated with dissipation of heat and with cathode emission increase with the power.

It is evident that the pulse operation used in radar and beacon equipment provides a perfect solution to these difficulties, wherever this type of operation is applicable; it permits high-voltage and high-current operation, because the low-duty cycle maintains the average dissipation and emission reasonably low. Where c-w operation, or high-duty cycle in general, is necessary, the problem is much more difficult.

CHAPTER 19

THE RESNATRON

By W. G. DOW AND H. W. WELCH

19-1. The Resnatron.—The *resnatron* is a high-power tetrode cavity resonator type of tube which, in terms of continuous power output at ultrahigh frequencies, has completely outdistanced any other development. In its present state, the resnatron is capable of delivering up to 60 kw of power at high efficiencies and at frequencies from 340 to 625 Mc. The theory of operation is similar to that of a Class C oscillator. Cavities comprising the oscillating circuits are within the vacuum envelope and are built integral with the electronic elements of the tube. For reasons both mechanical and electrical the tubes are continuously evacuated and demountable.

A schematic development of the resnatron is shown in Fig. 19-1. There are two resonant circuits consisting of coaxial cavities $\frac{3}{4}$ wavelength or $\frac{1}{2}$ wavelength long. One cavity, called the *cathode cavity*, is formed between cathode and control grid. Another cavity, called the *anode cavity*, concentric with and surrounding the cathode cavity, is formed between screen and anode. The actual electrode structures form parts of the cylindrical walls of the cavities and are located approximately at positions of voltage maximums on standing waves set up in the cavities in oscillation. Transit time is such that electrons accelerated in the cathode cavity give up their energy to the r-f field in the anode cavity. The sharp bunching required by a Class C amplifier or oscillator is not derived from velocity modulation, as in other u-h-f oscillators, but from direct formation of dense bunches of electrons by means of varying potential between the control grid and the cathode. Use of the screen grid or tetrode structure in the resnatron speeds up electrons in transit from cathode to control grid, thus dispensing with some of the transit-time problems experienced with u-h-f triode oscillators. By means of a small capacitive probe between anode and cathode, energy is fed back into the cathode circuit to maintain oscillation. The frequency of oscillation is determined by the electrical length of the cavities which, for tuning purposes, is variable.

Power is coupled inductively from the anode cavity through a tunable coaxial transformer into a copper waveguide transmission system. Plate efficiencies of over 70 per cent may be obtained under the proper condi-

tions. The high Q of the system makes wide-band modulation difficult, but with adjustment of feedback and loading, low- Q operation is possible with a very substantial reduction in efficiency.

A system built around the resnatron requires a 10- to 16-kv 100-kw power supply and a fast vacuum system capable of producing an ultimate vacuum of between 10^{-4} and 10^{-7} mm Hg pressure. Associated control circuits, modulating circuits, power-disposal systems, and maintenance equipment will depend on the use to which the resnatron is put.

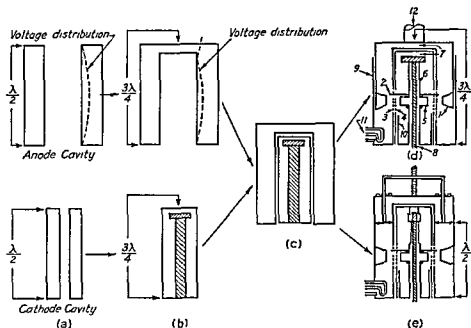


FIG. 19-1.—Schematic development of a resnatron from a pair of coaxial cavities. (1) Anode, (2) capacitive feedback probe, (3) screen grid, (4) control grid, (5) filament wire, (6) bellows for tuning, (7) tuning capacitance, (8) cathode-cavity tuning rod, (9) screen-anode by-pass condenser, (10) grid-cathode by-pass condenser, (11) coupling loop, (12) anode-cavity tuning rod.

Early work on tubes of the resnatron type was carried on in cyclotron laboratories, where the technique of producing large high-pumping-speed high-vacuum systems was developed. For mechanical and economical reasons, demountable tubes were necessary. The subsequent development of the resnatron, leading finally to truck-borne transmitter units, has indicated, under adverse conditions, the practicality of these techniques for producing high-frequency very high-power oscillations.

19-2. Structural Features of the Resnatron.—The schematic drawings in Fig. 19-1 show the essential components of a resnatron from an electrical point of view. Structurally these components are further complicated in form by cooling, mechanical, insulating, and vacuum require-

ments. A schematic cross section of a complete resnatron is shown in Fig. 19-2. Figure 19-3 is an exploded view of all the components ready for assembly, and Fig. 19-4 is a photograph of an assembled resnatron. This type of resnatron weighs about 250 lb. assembled.

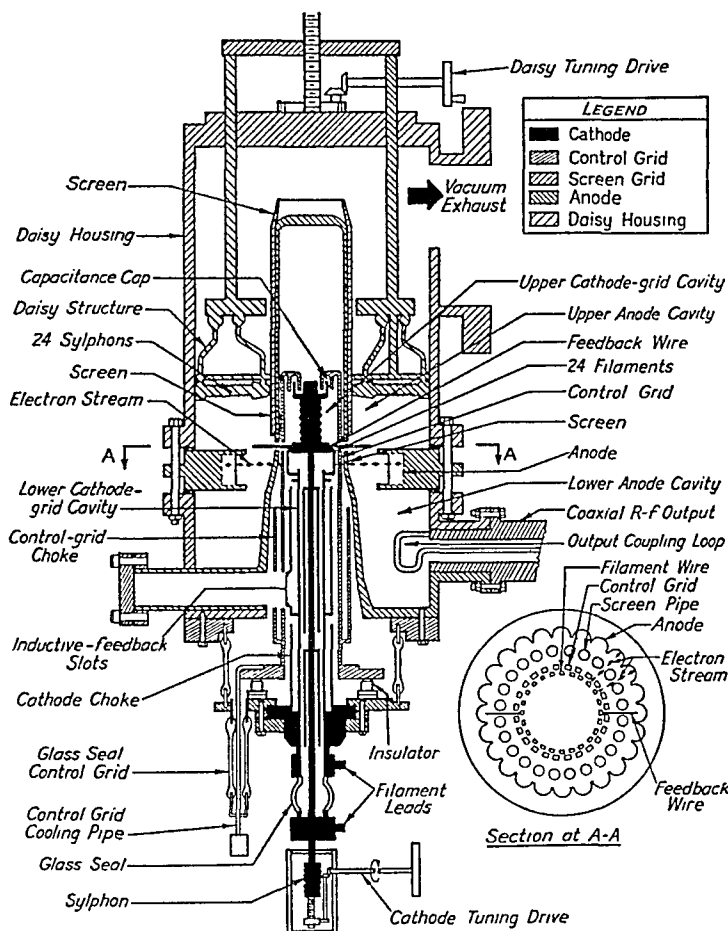


FIG. 19-2.—Vertical cross section of resnatron.

Figures 19-1 to 19-4 are self-explanatory and should be referred to frequently while reading the following sections, in which some of the structural features will be considered in more detail.

19-3. Cathode Cavity.—Cathode cavities of two types are shown in Fig. 19-5. Either type of cavity is electrically equivalent at the resonant frequency to a $\frac{3}{4}$ -wavelength transmission line short-circuited at the lower end and open-circuited at the upper end. Tuning of the cathode cavity is accomplished by moving up and down an internal rod, carried by metal bellows. This mounts at its upper end the cap spaced close to

the top of the cathode cavity to provide adjustment of the capacitive loading (see Fig. 19-5a).

The tuning capacitance, shown in Fig. 19-5b, is more sensitive to vertical motion of the tuning rod than the type shown in Fig. 19-5a and therefore more desirable where wide tuning range is the objective.

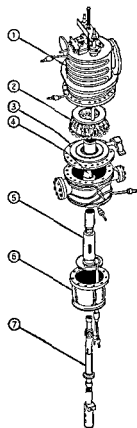


FIG. 19-3 — Isometric sketch of disassembled resonator (1) Tuner housing, (2) tuning plunger ("daisy"), (3) anode, (4) screen-grid assembly, (5) control-grid assembly, (6) cathode and grid insulator, (7) cathode assembly.

A choke-condenser by-pass and blocking arrangement is used to maintain r-f continuity of the cathode cavity where it is broken to provide d-c insulation between cathode and grid. This arrangement must serve two purposes. It must provide effective by-pass action, *i.e.*, appear like a short circuit from the interior, and it must prevent serious escape of r-f energy and appreciable coupling to other parts of the circuit. The ideal means of doing this is a coaxial section $\frac{1}{4}$ wavelength long terminated in a perfect open circuit, usually called a *quarter-wave by-pass condenser*. However, it is difficult to ensure a true open-circuit termination and quite impossible to achieve a $\frac{1}{4}$ wavelength for all frequencies in the tuning range of the resonators.

The formula for the input impedance Z , to a quarter-wave by-pass condenser of characteristic impedance Z_0 and terminated by an imperfect open circuit having an impedance Z_L is

$$Z = \frac{Z_0^2}{Z_L} \quad (19-1)$$

Obviously the smaller Z_0 the more nearly Z will approach a short circuit in spite of the fact that the termination does not have infinite impedance. To make Z_0 small, the two cylinders should be as close together as the d-c insulation requirements will permit. This is in accordance with the common-sense judgment that a good by-pass or blocking condenser is one with its plates as close together as possible consistent with successful blocking of the d-c voltage. However, this expedient is inadequate if the by-pass comes to be substantially more or less than $\frac{1}{4}$ wavelength long. It was therefore found necessary for complete coverage of the tuning range to lengthen the by-pass condenser in the shift from high- to low-frequency operation. This at the same time has the possibly more important effect of making the interior of the cavity longer. The cavity

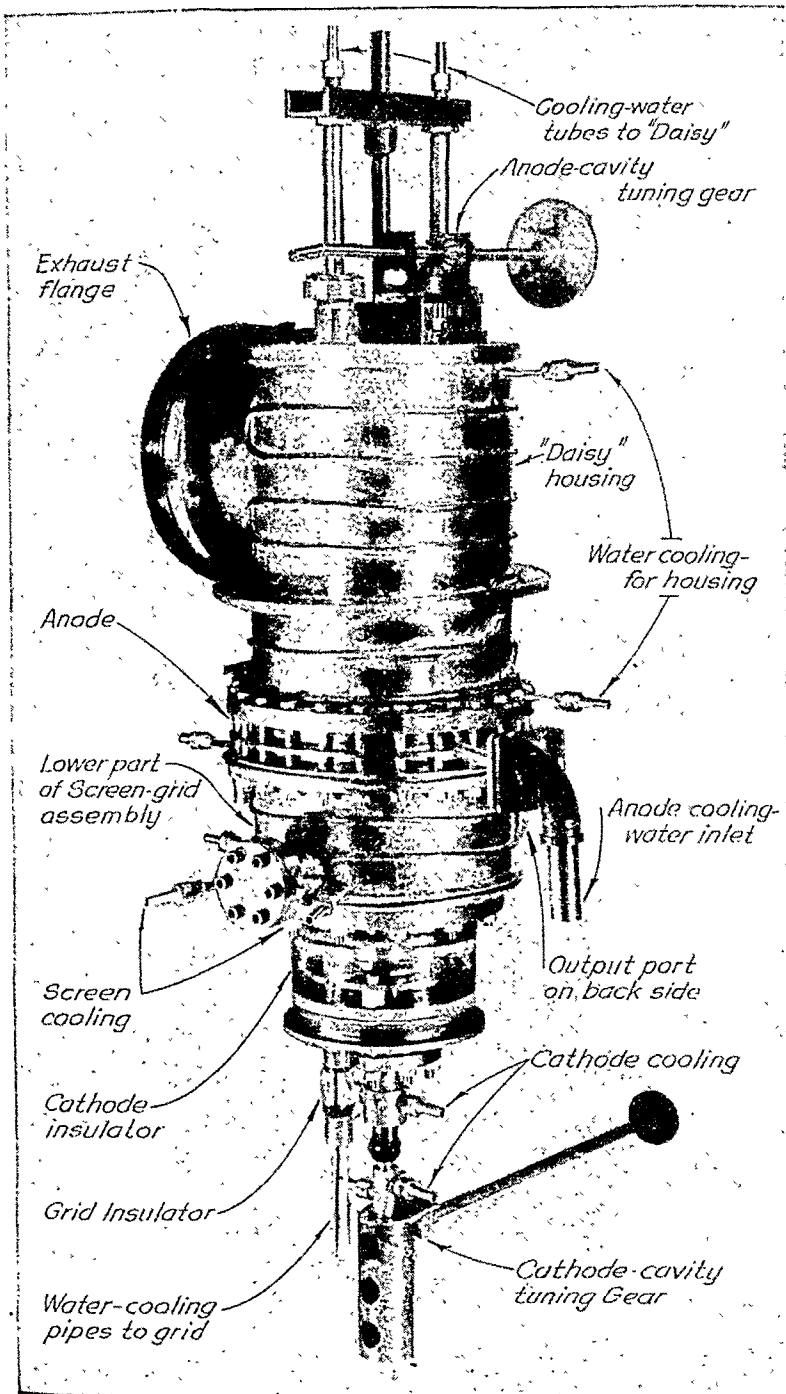


FIG. 19-4.—Assembled resnatron.

is also made longer at the upper end by raising the top caps of the grid and cathode structures which form the capacitive loading on the upper end of the cavity.

In operation, separate sets of grids and cathodes are used, one set for high frequency and one set for low frequency. In the case of the grids this is not absolutely necessary, since changing a grid from high to low frequency involves only a minor setscrew operation. In the case of the cathodes, however, the change is difficult, as it involves two solder joints

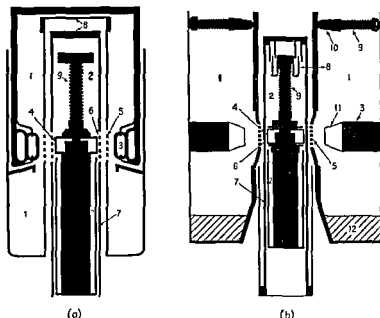


FIG. 19-5.—Vertical cross sections showing essential features of two types of anode and cathode cavities (1) Anode cavity, (2) cathode cavity, (3) anode, (4) filament wires, (5) screen grid, (6) control grid, (7) quarter-wave by-pass condenser, (8) tuning capacitance, (9) expanding bellows, (10) tuning plunger, (11) anode vane, (12) copper slug to shorten cavity.

one of which withholds cooling water at a pressure of 150 lb per sq in from the vacuum.

19-4. Anode Cavity.—Anode cavities of the type shown in Fig. 19-5a are built on much the same principles as the cathode cavities. With this type of anode, tuning is accomplished by pressure on the top of the cavity, which distorts the metal plate enough to vary the capacitance between the top of the anode housing and the screen cap. At 500 Mc the variation in frequency possible with this tuning scheme is about 30 Mc. The inner wall and the lower half of the cavity are at the d-c potential of the screen. The upper half of the outer wall and the top of the cavity are at anode potential. A by-pass condenser for the r-f current between screen

and anode structure is provided by extending the outer wall of the screen structure over the outer wall of the anode structure for $\frac{1}{4}$ wavelength. This presents a low impedance to the r-f current inside the cavity at the gap between screen and anode structures. This cavity, like the cathode cavities, may be thought of electrically as $\frac{3}{4}$ wavelength long with a short circuit at the lower end and capacitance loading at the upper end. The voltage distribution in this cavity is shown in Fig. 19-1b.

Figure 19-5b shows a form of the anode cavity that greatly extends the tuning range and, at the same time, reduces the Q of the cavity by eliminating energy storage in the tuning capacitance. The latter is desirable where wide modulation band width is the objective. The voltage distribution in this type of cavity is shown in Fig. 19-1a.

In this cavity the screen and anode are at the same d-c potential. The outer wall of the cavity is made continuous, and a sliding short-circuiting plunger having about 30-cm possible travel replaces the tuning condenser between screen wall and anode wall. The short circuit consists of a ring of sylphon bellows expanding radially between inner and outer walls of the anode cavity formed by screen structure and housing. This bellows ring is called the *daisy*, simply because it has the appearance of a daisy with the individual bellows as petals. The bellows carry considerable r-f current and are water-cooled. The water cooling (under 30 or 40 lb per sq in. pressure) also serves to expand the bellows, thus making good current contact with the walls. The daisy can be moved in the vacuum by means of rods passing through Wilson vacuum seals.

The daisy tuning plunger could easily accommodate the change in cavity length necessary to give the range of frequency desired (350 to 600 Mc). However, this change, about 2 to 1, would cause the electron stream to be badly displaced from the longitudinal center of the cavity in the low-frequency position. Such a design might introduce a tendency for higher harmonic moding (see Sec. 18-5). A better solution is to divide the range into two portions and, for the high-frequency portion, insert at the bottom end of the cavity a copper slug which effectively shortens the length of the lower half of the cavity. This slug is shown in place in Fig. 19-5b.

As indicated best in Fig. 19-2, the inner wall and lower portion of the anode cavity are formed by the screen structure, the upper part of the outer wall by the daisy housing, and the top by the daisy. The anode itself is mounted on the outer flange of the screen structure and supports the daisy housing. Both rubber and tin gaskets are provided at the joints at the top and bottom of the anode, for holding the vacuum and maintaining a current path. Extreme care must be taken in the approximately 100 soldered joints in the daisy assembly to avoid vacuum leaks. These leaks are especially troublesome because they are difficult to locate.

Vacuum leaks in other elements are not unusual but are in general easily located and repaired.

19-5. Electrode-structure

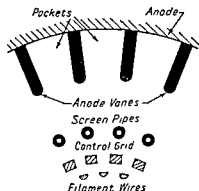


FIG. 19-6.—Horizontal quartered cross section of electrode structure.

Design. General Considerations.—The resnatron tubes are beam tetrodes employing fairly sharply focused electron beams, the beam section being a thin rectangle, as nearly as possible of the same size as the face of each filament, i.e., 0.85 in. long by 0.050 in. wide. One primary reason for designing the electrodes to give sharp focusing is to minimize the screen current, because otherwise, at the high screen voltages used, screen dissipation would be excessive. At the high voltage levels used in the resnatron, secondary emission,

usually a contributing factor to screen current in tetrodes, is not important. Another objective important in electrode design is to achieve the necessary focusing and plate-current values without drawing excessive grid current. Figure 19-6 is a sectional diagram of the electrode structure as used in the resnatron with the anode design that seemed to be the most satisfactory of those tried.

Since the resnatron design is based on the use of continuous evacuation techniques it is not only inconvenient but quite impractical to use a type of cathode requiring processing in a vacuum; in accordance with cyclotron demountable-tube practice, tungsten filaments are used as cathode surfaces. The 24 tungsten strips of the cathode "cage" are each 0.050 by 0.015 in. in section, the 0.050 in. face being the useful emitting surface. The useful length is $\frac{3}{4}$ in. The normal filament excitation is 1800 amp at 2.5 volts for a cathode with all filament strips intact. If one or two or more are broken but not short-circuited, the tube will still function, but at a reduced power level. Figure 19-7 shows a half-section cathode assembly. The cages can be removed and replaced by a simple soldering operation.

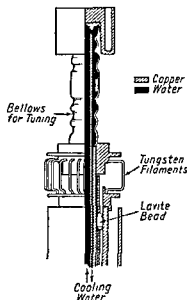


FIG. 19-7.—Section of cathode showing cooling-water channel.

Certain structural and cooling requirements affect electrode design.

The shape into which each of the 24 individual tungsten filament strips is bent provides an offset from the inner wall of the cathode cavity, so that this cavity can have appreciable radial depth, yet brings the emitting surfaces immediately out to a circle just inside the inner grid circle. It also permits the filament to expand lengthwise when heated, without serious sag or other displacement of its normal position relative to the grid.

The heating due to cathode-cavity circulating current on the interior surface of the grid cylinder, and that due to normal electron flow to the grid, is sufficient to make necessary special provisions for cooling the grid cylinder. The important path of heat flow away from the top half of the grid unit is via the grid cylinder to the built-in water channel shown in Fig. 19-8. In this figure the lower part of the grid is cut in section to show the water channel more clearly. It is desirable to keep the radius of the grid cylinder, especially where pierced by slits for electron transit, not too small to permit adequate heat transfer just below the slits.

The screen "wires" are hollow copper tubes, to permit the circulation of cooling water through them. These tubes extend inside the screen wall to the top of the structure, thus carrying away circulating-current losses as well as electron-current losses.

Attainment of Focusing Action.—Figure 19-9 gives the best pictured evidence of the actual sharpness of focusing of the beams. This is a

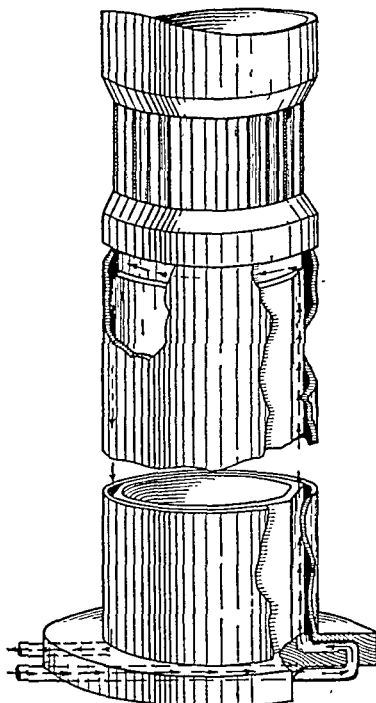


FIG. 19-8.—Section of control grid showing cooling-water channel.

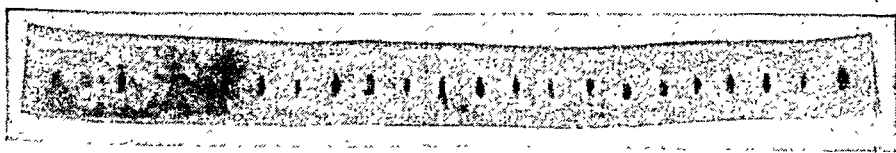


FIG. 19-9.—Photograph of copper anode strip showing effect of beam focusing.

photograph of a copper strip that was bent into the form of a circle, then inserted just inside the anode, forming in this way a temporary anode surface. A tube was operated for some time with this strip in position; the black markings are the results of electron-beam bombardment and indicate a very satisfactory sharpness of focus.

Sharpness of focus is obtained by attention to the details of the shapes of grid and screen sections relative to the flat strips of cathode surface, in order to obtain the proper electric-field configuration. Figure 19-10 shows shapes taken by equipotential surfaces in the region of the field where focusing is accomplished for the electrode voltages specified. Superposition of the two field patterns, Figs. 19-10a and b, properly weighted for voltage magnitudes gives the field for a particular pair of instantaneous voltages, if space charge is ignored. These equipotentials were obtained by electrolytic-tank measurements on scaled sections.

By calculations based on Figs. 19-10a and b or similar figures for other designs, it is possible to predict the electrostatic grid-to-screen μ -factor

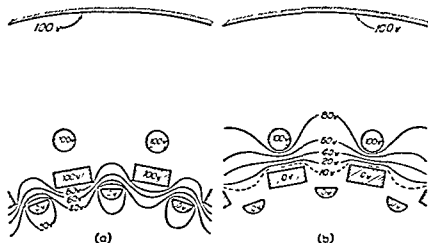


FIG. 19-10.—Equipotential maps around resonatron electrodes.

that gives the relative effectiveness of grid over screen in establishing an accelerating field at the cathode surface.

It is to be expected that focusing at ultrahigh frequencies will be somewhat different from that with d-c voltages because the fields will change during the flight of an electron. Probably the most important focusing action in resonatrons takes place while the electrons are still in the converging field in the grid-cathode pocket, because at that time electron velocities are low enough to give transverse fields time to act. In this region the r-f fields are strong, so that the focusing action during oscillation will almost certainly differ in detail from that with only d-c voltages applied. Focusing may be improved by the presence of the r-f field, or it may be made worse. The only significant experimental evidence is that the focusing into a thin flat beam is very satisfactory under conditions of normal use.

The details of the final electrode design are actually the result of a rather long trial-and-error process in which screen current and grid cur-

rent were important criteria of success. In the resnatrons with anode and screen at separate potentials, operation was considered satisfactory and normal with a screen current of around 300 to 500 ma when the plate current was between 4 and 5 amp. In tubes with no separation of the d-c potentials, it was impractical to measure screen current separately, but it was possible to judge from the rise of water temperature in the screen-cooling circuit that screen dissipation was not excessive.

Grid-screen μ -factor.—Aside from focusing the electrons into a beam, it is necessary in determining the proper electrode design to ensure that the screen makes the proper contribution to the electron-accelerating field at the cathode surface. This contribution depends primarily on two factors: the size and aspect ratio of the cathode-grid pocket and the angles involved and the radial positioning of the screen tubes.

If the cathode surface is made too wide in relation to the thickness of the grid wires, the effectiveness of the grid on electron emergence at the center of the filament surface is reduced undesirably. If the cathode surface is made too narrow in relation to the thickness of the grid wires, the grid-screen μ -factor becomes too large, and the screen does not aid sufficiently in electron acceleration at the cathode. Grid current is then likely to become excessive. Chamfer of the corners of the grid wires is a definite aid in keeping grid current low. Square outer edges will tend to collect an undesirably large number of electrons.

For any given screen potential the radial placement of the screen tubes is of some importance. If they are too far out, the cathode-surface electron-accelerating field of screen origin and the converging focusing field in the grid-cathode pocket are too weak. This will tend to cause excessive grid current for a given cathode current. By moving the screen tubes closer to the grid, the grid-to-screen μ -factor is decreased. This should permit drawing any given cathode current with a less extreme upswing of the grid and, therefore, less grid current. Since a larger negative swing is required for cutoff, the total grid drive is not seriously affected. The grid bias must be correspondingly increased to retain cutoff where it belongs in the cycle in order to keep the angle of plate-current flow small.

There are at least two limits to any desirable effects that may be achieved by moving the screen in. One is the ultimate requirement, in the face of a falling grid current, of an excessively large grid bias in relation to the size of the self-bias resistor. Another is that as screen tubes are moved in they begin to be electrostatically hidden behind the control-grid sections, and the effect of the screen on the cathode gradient therefore ceases to increase with reduction of screen radius.

19-6. Anodes.—Anode structures are illustrated in Fig. 19-11. In the anodes shown in Figs. 19-11a and b, the actual anode surface is formed

by pipes carrying the cooling water. In models shown in Figs. 19-11c and d, a single water channel behind the anode surfaces serves to carry away losses. The final design as shown in Fig. 19-11d was a compromise of several objectives. This anode is very simple to construct, being cast and machined.

The flow of water and the thickness of the copper wall between the electron beam and the water is such as to permit sufficiently rapid removal of heat. The final design, with 20 gal per min of cooling water at 180 lb pressure, has dissipated as much as 100 kw in static tests. The type

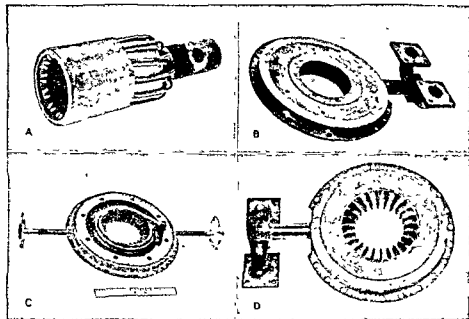


FIG. 19-11.—Photographs of different types of anode structures.

shown in Fig. 19-11b proved incapable of handling large amounts of power.

The "pockets" in Fig. 19-11d have two side walls and a back wall but no top or bottom walls. These pockets were to act as a trap for secondary electrons by making most of the secondaries strike walls of the pockets rather than reach the screen. This feature became less important at the high voltages used, as shown by comparison of the tetrode characteristics for low screen voltages, 100 to 500 volts in Figs. 19-13b. The sharp rise in plate current that occurs just as the plate voltage passes from below screen potential to above screen potential is a good approximate measure of the number of secondaries emitted by the anode when it is at screen potential, due to primaries whose energies correspond to screen potential. The ratio of secondaries to primaries in the tetrode characteristics in Fig.

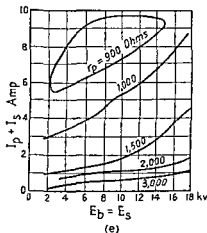
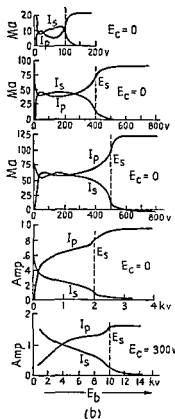
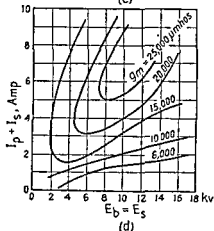
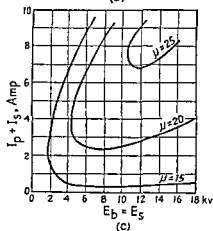
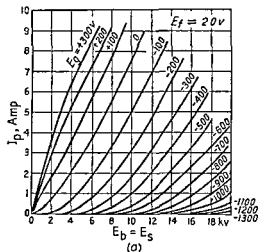


FIG. 19-13.—Some static characteristics of the resonatron.

One of the significant figures obtained from these static curves is the grid-to-screen μ -factor. This lies between 15 and 25, depending on the details of screen placement and grid design mentioned. Also of interest are the values of transconductance. These lie between 6000 and 25,000 μmhos , depending on the values of screen and grid voltages used, as well as on the details of screen placement and grid design.

19-8. Equivalent Circuit.—The structural features of resnatron construction have been described in the preceding sections. Two resonant coaxial cavities, one within the other, are built integral with the electrode structure of a beam tetrode. These have been described as equivalent to $\frac{1}{2}$ - and $\frac{3}{4}$ -wavelength cavities at the resonant frequency of operation.

The analysis of resnatron behavior can be simplified to lumped-constant circuit analysis by use of the equivalent circuit shown in Fig. 19-14. A more exact equivalent can be given by showing the various by-pass capacitors and the tuning capacitors. For simplicity, however, we shall assume that by-pass action is effective and that tuning can be accomplished in both resonant circuits by one means or another. Thus we may represent these capacitive elements as short circuits or as lumped with the existing capacitance. The circuit shown is usually known as the *grid-separation* or *grounded-grid circuit*. If the feedback capacitance is omitted, use as a tuned u-h-f amplifier is possible.

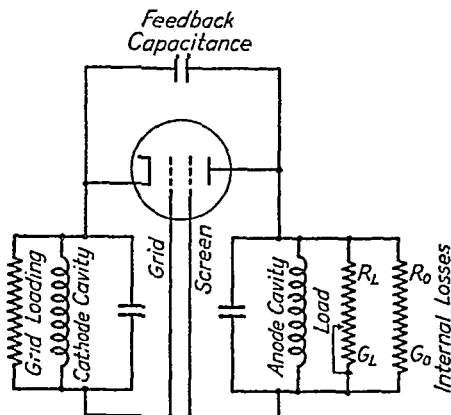


FIG. 19-14.—Equivalent circuit of resnatron for ultrahigh frequency.

The output loading is indicated in Fig. 19-14 as being adjustable. The mechanism of adjustment is contained in the output coaxial line-to-waveguide transformer. Internal copper-loss-loading conductance is indicated separately in the diagram, but may be thought of as lumped in with the output-loading conductance. The grid loading actually includes power losses of the following four distinct kinds:

1. Circuit losses in the cathode cavity
2. Power dissipated in the grid self-bias resistor
3. Grid dissipation due to electrons striking the grid
4. Cathode dissipation due to cathode back-heating, i.e., electrons returning to the cathode after receiving energy from the r-f field during their time of flight

The feedback capacitance consists of a probe extending from the cathode through the grids into the anode cavity. For the present it

will be assumed that feedback is in the proper phase and magnitude to maintain oscillations. A more detailed analysis of the feedback requirements will be given in Sec. 19-13.

19-9. Q of the Anode Cavity.—A simple analysis was given in Chap. 18 to show the nature of the relation among circuit constants, shunt impedance, and Q of the anode tank circuit for normal operation of a circuit of the type shown in Fig. 19-14.

Although the anode cavity shown in Fig. 19-5a is in a sense a distributed-constant circuit, a study of the figure suggests that in fact it is more nearly a lumped-constant circuit. The anode and the r-f shielding ring just below it intrude into the cavity sufficiently far to be the controlling factors in the cavity capacitance; the tuning capacitance at the top has to be fairly prominent in order to provide adequate tuning; thus, there are two lumped capacitances in series, while the circuit is completed across the closed bottom end of the anode cavity. Thus, the lumped-constant type of analysis given in Chap. 18 is probably as nearly correct as a distributed-constant analysis based upon a uniform $\frac{3}{4}$ -wavelength coaxial cavity would be. From either analysis, it is apparent that the loaded Q could be reduced by taking the following steps:

1. Enlarging the outer radius of the anode cavity, thus increasing the size of the current loop, and raising the value of L .

2. Changing to a plunger-tuned half-wave cavity instead of the capacitance-tuned three-quarter-wave cavity, thereby eliminating the capacitive energy storage in the tuning capacitor, and correspondingly lowering the effective value of C .

3. Reducing to an absolute minimum the capacitance from anode to screen in the neighborhood of the anode. This would be aided by eliminating the shield ring beneath the anode and by refinements in anode-design details.

All these changes are incorporated in the cavities shown in Fig. 19-5b. The enlargement of the outer radius is carried as far as is practical without giving the tube excessive weight and bulk. The daisy, as described in Sec. 19-4, is used to furnish plunger-type tuning. The anodes are designed with screen-anode capacitance in mind as well as other considerations.

19-10. Shunt-impedance Requirements.—In the discussion of shunt impedance in Chap. 18 the first three of the following considerations were mentioned:

1. From the standpoint of efficient operation the shunt impedance should not be below the optimum value, but can be moderately above it without marked ill effects.

2. The optimum shunt impedance is roughly proportional to the d-c plate supply voltage.

3. For a given angle of operation, the optimum shunt impedance is roughly inversely proportional to the d-c plate current.

4. In wide-band operation, the loaded Q , and therefore the shunt impedance, should not exceed a given value.

Thus 1 and 4 in combination indicate that the proper shunt impedance to use, where low- Q wide-band operation is wanted, is just the optimum value, no more, no less, while 2 and 3 in combination indicate that, in order to obtain the low shunt impedance needed for low- Q operation, tubes should be used that operate satisfactorily at as low d-c plate voltage and as high d-c plate current as possible. These considerations dictated the changes from the type of anode cavity shown in Fig. 19-5a to the type shown in Fig. 19-5b. The Q is lowered by decreasing C and increasing L . It is also possible by modifications in output coupling to reduce markedly the tank-circuit shunt impedance which in turn allows operation at lower plate voltages and higher plate currents. For high shunt impedance (high Q), operation with 16,000 volts on the plate at 2.5 to 3 amp is typical. By reducing load impedance (low- Q operation) these values are changed to 10,000 volts at 4 or 5 amp.

It is possible to provide feedback coupling adequate to ensure even with moderate r-f plate voltages the large grid drive necessary to draw the large currents required for low- Q operation.

19-11. Parasitic Oscillations.—Almost any oscillator while under development will have unwanted modes of oscillation, and the resnatron is no exception. Several types of parasitic resonances have been observed. One was an accidental 1000-Mc resonance in the cavity between control grid and screen grid. This was eliminated by adding capacitance between the top of the control grid and the top of the screen grid. Another resonance at almost the same frequency was attributed to a circumferential mode excited by asymmetrical placing of the feedback wires. A third resonance, obtained only in the type of cavity shown in Fig. 19-5a, is attributable to the $\lambda/4$ mode, possible because of the open circuit at the top of the cavity. This mode is especially likely to occur when the desired $3\lambda/4$ mode is being heavily loaded for low- Q operation. In this case the $\lambda/4$ mode, which is normally not heavily loaded, probably has a higher Q than the desired mode. The tendency of the oscillating system would be of course to oscillate in the high- Q rather than the low- Q mode. This parasite may be eliminated by a small series-resonant circuit coupled into the cavity, which detunes the undesired mode until feedback conditions necessary to excite the mode are no longer satisfied.

19-12. Transit-time Effects.—The effect of transit time on the operation of triodes and tetrodes at ultrahigh frequencies has been discussed in Chap. 18. The transit angle for the time of flight of electrons from the

cathode to the grid is a large fraction of the total transit angle in the resnatron, because when the electrons first leave the cathode they are moving very slowly. The electron stream does not constitute plate current until it passes the screen, so the grid-to-screen transit angle adds to this effect. Also, plate current exists as long as any electron remains in transit to the anode, so that perhaps half of the screen-to-anode transit angle also contributes to the total transit-angle delay between grid voltage and fundamental frequency component of plate current. This angle must be added to the 180-deg phase difference between grid voltage and plate voltage necessary for optimum performance in ordinary low-frequency oscillators of this type. The effect of transit angle on feedback is discussed in Sec. 18-4.

No accurate measurements of transit-time phase delay were made in the resnatron, but rough calculations indicate that phase delays between 30 and 120 deg should be expected, depending on details of handling, voltages, spacings, etc.

In triodes a considerable difference in velocities of the individual electrons entering the grid-anode space results from the fact that all electrons do not leave the cathode at the same time during a cycle of oscillation. The introduction of a screen between the grid and the anode in the resnatron causes *all* the electrons that traverse the grid to be rapidly accelerated to such an extent that the differences between the grid-traverse velocities become quite unimportant as far as any effect on the time it takes the electrons to reach the plate is concerned. Thus the introduction of the screen suppresses the debunching action incidental to cathode-to-grid transit, and so tends to permit maintaining a good operating angle, if other conditions are right.

The introduction of the screen grid also reduces the time of flight of the electron through the interaction region between grid and anode in which the electron by its movement converts d-c power into a-c power. By reducing this *interaction time* a large operating angle, which results in poor efficiencies, is avoided. (Plate current associated with any electrons starts flowing when the electron passes the screen and continues to flow until the electron reaches the anode.)

Cathode back-heating caused by the return to the cathode of electrons which left late in the cycle is fairly marked in the resnatron. It shows up as an increase in filament resistance so that as power output is raised filament current falls off. It is necessary to reduce the filament heating power by between 15 and 40 per cent during operation, if it is desired to keep the filaments at their normal rated temperature and therefore resistance. Cathode life is shortened if this is not done. The resnatron cathode is tungsten and is designed for large heating power. In thoriated

or oxide-coated filaments requiring small heating power, back-heating can cause immediate destruction of the filament.

19-13. Feedback Circuit.—A qualitative analysis of feedback conditions in the resnatron is all that can be given because of the many unknown factors in resnatron operation.

Two feedback schemes were used with the resnatron. These are shown in the equivalent circuit in Fig. 19-15. Points P , G , and C refer, respectively, to the plate, grid, and cathode. It will be noted that the heavy line portion of this circuit is essentially the same as Fig. 19-14 except for the substitution of the equivalent current generator $y_m E_g$ and the plate admittance of the tube y_p for the electrode structure shown in Fig. 19-14.

The dotted lines indicate the first type of feedback used on the resnatron, known as *external feedback* because the feedback power is

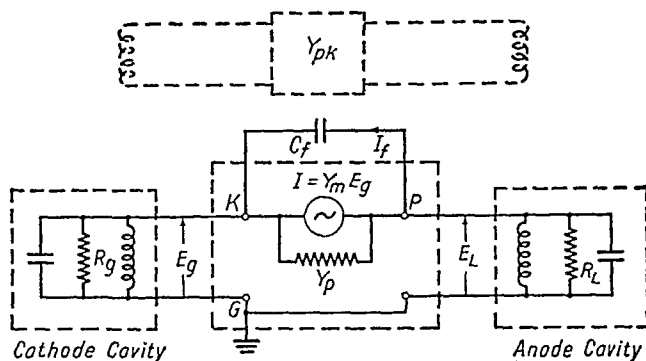


FIG. 19-15.—Equivalent circuit of resnatron for analysis of feedback conditions.

actually coupled out of the anode circuit, carried through a coaxial line containing tuning stubs and back into the cathode cavity. This method is theoretically desirable because the section of transmission line makes possible variation of phase of feedback as well as magnitude. However, the additional vacuum joints, glass output seals, tuning plungers, etc., are the source of serious practical difficulties, which led to the use of an *internal feedback* system consisting of molybdenum wires fastened to the cathode and extending through the grids to the anode. These wires form a capacitance between the anode cavity and the cathode cavity. This capacitance is shown as C_f on the diagram. This type of feedback has the disadvantage of being adjustable neither in magnitude nor in phase. It does have the advantage of simplicity in construction and can be made to work by proper tuning of the anode and cathode cavities.

The circuits shown in Fig. 19-15 are variants of the general equivalent network of a grounded-grid oscillator discussed in Secs. 18-8 to 18-10

and shown in Fig. 18-4. It is unnecessary to repeat the discussion of the conditions for oscillation and the phase relations obtained during steady-state oscillation since presentation in Chap. 18 and the vector diagram of Fig. 18-9b apply directly to the resnatron oscillator.

It should be pointed out that the Q 's of the cavities are important since in a resonant circuit the rate of change of reactance with frequency at resonance is proportional to the Q of the circuit. This indicates that if the cathode cavity is a lower Q circuit than the anode cavity it will have less effect on the resonant frequency. The smaller effect, which would be the normal expectation, is actually observed in operation.

On the other hand, as has been shown above, the tuning of the cathode cavity will determine the phase of the feedback voltage and thus the efficiency. This also was observed in operation.

Some recent tests have been made with an adjustable feedback capacitance mounted on the anode instead of the cathode and controlled through Wilson vacuum seals. This proved quite helpful for wide-tuning-range operation.

19-14. Output Coupling.—An important problem in the use of the resnatron is to provide an output system that will load the tube with the proper shunt impedance and provide a sufficiently rugged vacuum-tight avenue for the r-f power to pass from the evacuated space to an atmospheric transmission line.

The system finally used with the resnatron is shown in schematic cross section in Fig. 19-16. A loop in the lower part of the anode cavity couples power into a coaxial transmission line. Loops of two sizes were used, one about 2 by 1 in., for the high-frequency range, and one about 2 in. square for the low-frequency range.

The coaxial line-to-guide transformation is accomplished by passing the coaxial line completely through the short (6-in.) dimension of the guide. Inside the guide the outer conductor is replaced by a glass cylinder that forms the vacuum seal through which power is radiated by the inner conductor. The inner conductor thus acts as an antenna across the narrow dimension of the guide. The entire interior of the line is evacuated through the opening into the anode cavity.

The reactance of the loop is compensated for, and optimum phasing of currents in the inner conductor serving as an antenna is obtained, by means of a tuning plunger in the section of the coaxial line extending through the guide. The length L is made more than $\frac{1}{2}$ wavelength for any frequency of operation so that any phasing desired can be obtained. This plunger is operated through Wilson vacuum seals.

The coaxial output line is passed through holes in the guide that are near one side of the guide, as shown in Fig. 19-16b. Because the impedance of the guide increases from the edge to the center, by choice of a

correct position for these holes a fairly satisfactory broad-band match can be obtained.

In early work with the resnatron this general type of transformation was used but with a movable plunger in the stub end of the waveguide below the passage of the coaxial line. It was found, however, that this tuner and the coaxial-line tuner were so close together that their actions were not independent, and the two adjustments gave little greater flexibility than the coaxial-line tuner alone. The waveguide stub was therefore closed off with a copper plate about $\frac{1}{4}$ guide wavelength from the coaxial-line passage.

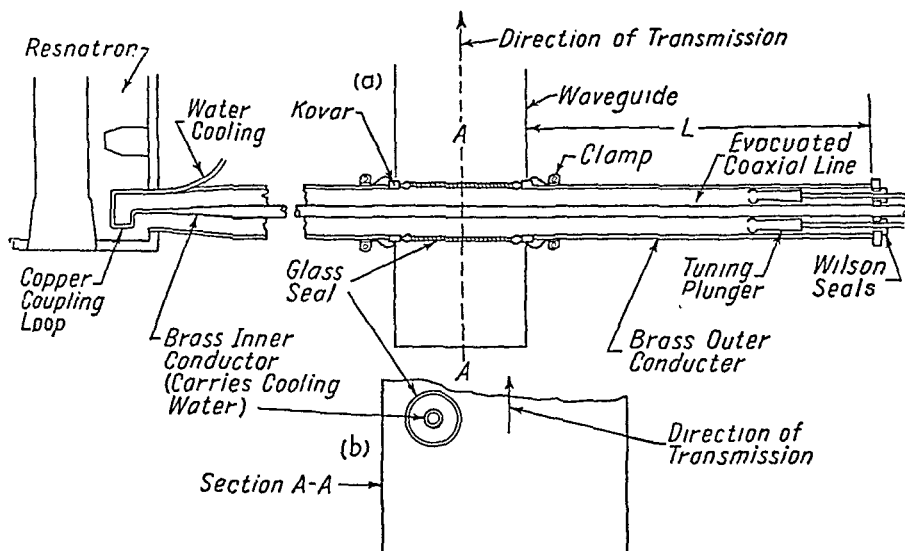


FIG. 19-16.—Schematic drawing of output coupling system.

The loops and the inner conductor are made of tubing to permit passage of cooling water. The loops are of copper, but the coaxial line is of brass to improve the functioning of the sliding contacts of the tuning plunger. A Kovar-to-glass seal is used in the glass cylinder. The section containing the glass is soft soldered in place and can easily be replaced in case of breakage.

In Fig. 19-15 the load impedance R_L is shown as a pure resistance. Actually, of course, it will have a reactive component and will therefore affect the frequency of oscillation. At every operating frequency there is some position of the output tuner that makes the output circuit resonant. At this position the shunt impedance is low and, depending on the mutual inductance of the coupling loop, may be lower than optimum for the given operating conditions. If this is the case, optimum loading occurs when the tuner is slightly detuned from this resonant position. Since it can be detuned in two directions, there are two set-

tings for optimum load impedance, but they differ in that the reactances coupled into the load are of opposite signs and therefore result in operating frequencies that may differ as much as 10 Mc. This means that, when the tube is oscillating with feedback prongs providing grid excitation adequate to maintain oscillation over the entire range of output tuner adjustment, the grid excitation for maximum drive, the plate-circuit efficiency, and the frequency will vary through a characteristic cycle as the tuner is adjusted. This effect is shown in the data presented in Fig. 19-17.

19-15. Transmission Line.—In the use of the resnatron transmitters it was necessary to locate the antennas between 100 and 200 ft from the transmitters. For transmission of 15 to 30 kw of power over this distance

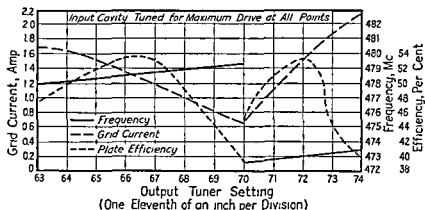


Fig. 19-17.—Behavior of resnatron with changes in loading.

coaxial line seemed impractical because of loss and mechanical fabrication difficulties. This led to the choice of waveguide.

In order to handle the power adequately at both extremes of the tuning range two sizes of guide were used, one for high-frequency tubes and one for low-frequency tubes. For high-frequency operation the guide used was 6 by 15 in.; for low frequency the guide was 6 by 22 in.

It was found by experiment that simple bolted flange connections with copper facings bent back from the main length served as entirely satisfactory joints between sections of guide. These were used throughout and ensured greater freedom from frequency sensitivity than could have been obtained with other radically different types of joints.

19-16. Power Disposal.—Two distinct types of load-disposal devices were needed in experimental work with the resnatron: one to serve either as a wattmeter or as a calorimeter to calibrate various types of wattmeter; the other to serve as a simple means of power disposal without much reference to power measurement. It was necessary for the first type to have quick response and get into thermal equilibrium in a minute

or two for any given condition, which required using a small amount of water and passing it through very rapidly. However, it was permissible for this device to be frequency sensitive, as tuning devices were available for matching it to the line at various frequencies. The second type of device needed to offer a flat termination for all frequencies, but quickness of response or measuring provisions were unimportant. Both types have been devised and are described in Chap. 24.

19-17. Operating Data.—With c-w operation, operating plate-circuit efficiencies from 40 to 60 per cent at 50- to 75-kw output were observed with the tubes that had the best type of anode. Up to 70 per cent at 30-kw output was obtained with tubes of an earlier design having narrower tuning range and higher Q . The actual efficiency of interaction was greater than these figures, because cathode and grid losses had to be supplied from the tank circuit.

Table 19-1 gives typical values of electrode voltages and currents used in c-w operation of the resnatrons (see Fig. 19-18). The data are for a low-frequency resnatron with an anode of the type pictured in Fig. 19-5b, equipped with three feedback wires each extending $\frac{1}{2}$ in. beyond the screen pipes, with filament at 2.0 volts, an average operating vacuum of less than 10^{-6} mm Hg, and a 2- by 2-in. output loop. Grid excitation was tuned for maximum plate efficiency, and in a direction to make the input cavity look capacitive to the feedback voltage.

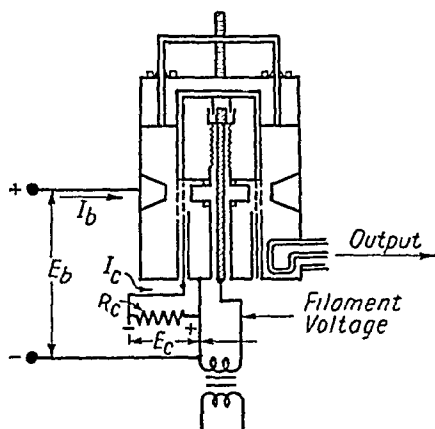


FIG. 19-18.—Direct voltages and currents in resnatron circuit.

TABLE 19-1

λ , cm	E_b , kw	I_o , amp	kw_{in}	kw_{out}	Plate efficiency, %	E_c , volts	I_c , amp	R_c , ohms
75.0	13.0	6.2	80.6	50	62.0	1300	1.20	1084
75.0	15.5	7.2	111.0	70	63.1	1900	1.15	1650
75.0	16.7	7.2	120.0	75	62.5	2000	1.00	2000
75.0	16.4	7.9	129.4	80	61.8	1600	1.40	1140
75.0	17.0	8.1	137.5	85	61.8	1500	1.40	1070

19-18. Modulation.—In order to meet the requirements for use it was desirable to be able to modulate the resnatrons with modulating fre-

quencies from about 50 kc to about 2.5 Mc. Because of this broad frequency range, video amplifiers were necessary in the modulation circuits, and for that reason the modulation frequencies, power levels, etc., will be spoken of here as the video frequencies, video power levels, etc.

At the 2.5-Mc end of the video-frequency range, the chief problem is to obtain resnatron r-f circuits with a Q low enough to give the 5-Mc band width necessary to accommodate the 2.5-Mc side bands. A reasonably satisfactory Q would provide 4 Mc between half-power points. By using methods discussed earlier in connection with loading and feedback, it was found possible to approach this band width under some conditions, although at a sacrifice in efficiency.

This type of modulation can be produced by swinging the grid voltage, or the plate and screen voltages simultaneously, at a video rate. The video-frequency modulating power required of the output stage of the modulator is proportional to (1) the upper video-frequency limit, (2) the capacitance to ground of the resnatron electrode driven by the modulating voltage, and (3) the square of the voltage required for modulation.

The reason for this is of course that the effective video-frequency conductance to ground of the driven electrode must be made high enough to give, in connection with the capacitance burden, a video-output-stage time constant (RC) small in inverse proportion to the upper video-frequency limit. A noninductive shunting resistor must be added externally to get this conductance, if the internal video-frequency conductance to the driven electrode is not sufficiently high. Such a shunting resistor was used in the resnatron transmitter circuit. The video power required in the output stage for any given total value of output-stage conductance is of course proportional to that conductance and to the square of the video voltage that is required to accomplish modulation.

The capacitance from the grid or from the plate and screen to the cathode amounts to a few hundred micromicrofarads. The modulation voltage necessary to drive the anode-screen combination must have a peak value approaching the anode d-c voltage, i.e., between 8000 and 10,000 volts. It is possible, however, to get reasonably effective grid modulation with a video voltage having a peak value amounting to between 2000 and 3000 volts.

Calculations indicated that it would require about 5 kw of video-frequency power to provide grid modulation and at least 15 kw to provide plate-and-screen modulation. Grid modulation was therefore employed as requiring less equipment and less power. Even to get 5 kw of video power required putting about 15 kw of d-c power into the plate circuit of the modulator tubes.

The last stages of the modulating circuit used in the completed trans-

mitters employed the circuits shown in Fig. 19-19. The final stage used Federal Telephone and Radio type-125A water-cooled tubes, operating not quite Class A. Since the grids of the 125A tubes offer purely capacitive loads to the driver stages, it is possible, with the circuit shown, to drive either one or two 125A tubes from the same driver stage. Thus one 125A tube was provided to modulate each resnatron, but one modulator driver stage could drive two 125A's.

It is not usually possible to obtain satisfactory grid modulation of an oscillator, for reasons having to do with the biasing circuit. In order to be self-starting, an oscillator must employ self-bias rather than fixed bias, and any simple self-biasing circuit tends to suppress changes in grid-bias

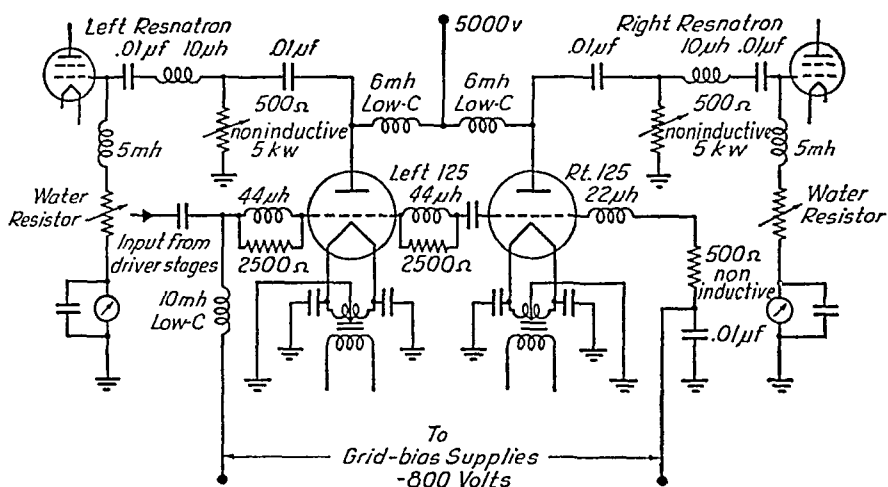


FIG. 19-19.—Resnatron modulator, last stage.

voltage from such causes as change in biasing resistor or introduction of a d-c voltage in addition to the bias. Thus to permit grid modulation at all, the biasing circuit must be given a time constant long compared with modulation-frequency periods. On the other hand, satisfactory oscillator stability, as, for example, when plate modulation is used, is best if the grid-bias-circuit time constant is short compared with modulation periods. Also, it is difficult to approach 100 per cent grid modulation closely except by using relatively little or even no grid bias. These problems are familiar in the radio art and involve no new principles. It was, however, found possible to achieve a combination of circuit constants that gave about 80-per cent grid modulation with satisfactory stability for the video band width used. Linearity was not good, but that was not important.

Under certain conditions of operation the resnatrons exhibited an inverted type of grid modulation, in which oscillation died out at the maximum *positive* swing of the grid modulation voltage and was a maxi-

mum at the *negative* swing. This was observed at various high and low modulation frequencies, and was studied briefly but with some care at a 400-cycle modulation frequency, using oscilloscopes to verify polarities, rise and decline of oscillation, etc. At times both normal and inverted modulation response were observed simultaneously, showing fading out of oscillation at both peaks and valleys of the modulation-voltage cycle. The study of the modulation of the final transmitter was not carried out in sufficient detail to enable any assured determination as to which of the two types of modulation predominated. Usable modulation was obtained, which was, for wartime purposes, the important consideration. Reasonably stable behavior of either type of modulation was obtainable in the laboratory.

The percentage of modulation was limited by a tendency of the tube to go out of oscillation, and sometimes remain out, if the modulation swing was enlarged to carry the modulation too close to 100 per cent. It might then be necessary to remove the modulation voltage in order to start oscillation again. The closeness of approach to 100 per cent modulation attainable was sensitive to a variety of adjustments, including the Q of the resnatron circuit, the degree of detuning of the cathode cavity, the amount of feedback used, and the grid-bias resistor. Achieving maximum modulation effectiveness was for a time more of an art than a science, but means were devised for monitoring which permitted reasonably good standardization of techniques, although a satisfactory scientific understanding of the reasons for the interplay of the various controls was never achieved.

19-19. Resnatron Transmitters.—The purpose of this chapter has been to outline the features of the resnatron oscillator and to mention methods of analysis, techniques, and concepts that are useful in working with the resnatron. Very little has been said about associated equipment required for a complete transmitter.

The resnatron tubes formed the basis for truck-mounted transmitter units to be used in the field and therefore to be self-contained. To carry the components for each transmitter unit the following 10 trucks (eight 10-ton) and a trailer were used:

Three primary power-generation trucks, each carrying a 75-kw diesel engine-generator unit with generator switching and synchronizing equipment

Two rectifier trucks, containing transformers and tube racks for the main power rectifier, which supplied 100 kw of d-c power, and for the modulator rectifier, which supplied 35 kw of d-c power; also remote-control switching, protective, and relaying apparatus

Two transmitter trucks, each mounting two resnatrons with associated output systems, vacuum pumps, water-circulating system, heat exchanger and

blower, modulator, panels for determination and control of vacuum, and panels for control of resnatrons and modulator

Two antenna trucks

One workshop truck

One water trailer

The various components were selected to permit simultaneous operation of any two of the four resnatron oscillators mounted in the trans-

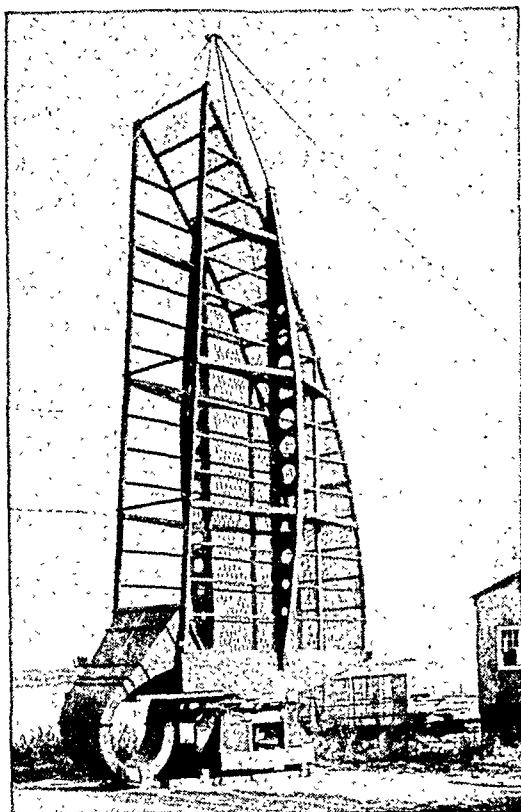


FIG. 19-20.—A swivel-mounted horn-fed antenna of the parabolic half-cheese type.

mitter trucks, the other two tubes normally being kept under vacuum in stand-by position ready for immediate use. One Federal Telephone and Radio type 125A modulating tube was mounted with each of the four resnatrons, because of the electrical necessity for close positioning of the modulator-output stages to the resnatrons. Each transmitter truck carried low-level and driver modulator stages capable of driving any two of the four 125A tubes.

The resnatrons delivered power through a short tunable coaxial-type transformer into a copper waveguide transmission system. Enough waveguide, with necessary special measuring and switching components,

was provided to serve two antennas each located some 200 ft from the transmitter trucks.

One tube in each truck was normally assembled for use in the low-frequency band and connected to a 22-in. guide system, the other tube being assembled for high-frequency operation and connected to a 15-in. guide system. Each antenna truck carried a swivel-mounted horn-fed antenna of the parabolic half-cheese type (see Fig. 19-20).

Although the two antennas were alike as to reflector size and dimensions, one of them was equipped with a horn designed for 460- to 625-Mc operation and connected to the 15-in. guide transmission system, while the other had a 340- to 520-Mc horn served from the 22-in. guide system.

The vacuum system for resnatron use must be a high-speed system in order to pump out quickly gas formed from arcs in the tubes and to enable internal changes to be made quickly in emergencies.

The continuously pumped-tube operation worked out more smoothly in field service of the truck-mounted assembly than the most optimistic predictions had anticipated. The use of the continuous evacuation system proved to be no more difficult in the field, after establishment of routine, than in the laboratory. It was found entirely practicable to train a crew of radio technicians, having no formal education above the elementary school level, to maintain the vacuum system in excellent operating condition, to learn how to locate and correct leaks, make minor tube repairs, do the soldering involved at certain points in assembly, etc. Furthermore in field use the vacuum-system conditions became increasingly better with passage of time because of the improved cleanliness of the vacuum system resulting from hours of continuous operation. It was possible for the crew to replace used cathodes and change feed-back prongs, a change that involved breaking and reestablishing the vacuum, without impairing the general high-quality level of vacuum conditions obtained as a result of continuous use.

An indication of the practicability for field use of this type of apparatus can be obtained from the time required for an experienced crew to get the outfit into operation in a new location. This involves unpacking or uncrating the tube parts, assembling the tubes, mounting the 125A tubes in their water jackets, checking the operation of all the control, relaying, and water-circulating systems, assembling the waveguide and r-f transmission system, establishing the vacuum, outgassing the tubes, and getting full power adjustments made. On one occasion, when an accurate check was made, this was accomplished by an experienced crew of eight men in 45 working hours.

CHAPTER 20

PRINCIPLES OF MAGNETRON OPERATION

BY GUNNAR HOK AND W. G. DOW

20-1. Why and How Continuous-wave Magnetrons Were Developed.

To the general trend of power generators toward successively lower efficiencies at successively higher frequencies, the magnetron is a remarkable exception. Efficiencies up to 70 per cent have been measured at frequencies as high as 3000 Mc and of the order of 50 to 60 per cent up to 10,000 Mc. After this fact had been established, naturally the development of power generators for frequencies above 1000 Mc concentrated on magnetrons. For special applications, magnetrons have also been used to advantage in the range 200 to 1000 Mc. The magnetron oscillators thus deserve consideration in some detail, at least until a superior combination, comparable to the crystal-controlled master-oscillator-power-amplifier transmitters of lower frequencies, has been developed.

As far as c-w magnetrons were concerned, an extensive development program was mapped out on paper during the war emergency, calling for development of broadly tuned magnetrons at a number of discrete power levels, approximately 50 watts, 150 watts, 1 kw, and 10 kw, to span a frequency range of 100 to 11,000 Mc. Complete frequency coverage at all these power levels was not, however, intended. For frequency coverage, the emphasis was placed on the 150-watt and 1-kw levels, the former for 100 to 1200 Mc, the latter for 500 to 11,000 Mc. The function in this program assigned to the Radio Research Laboratory was to assist in the technical guidance and in the evaluation of the tubes, particularly with reference to the coordination of tube and equipment design.

20-2. Relationship between Continuous-wave and Pulsed Magnetrons.—At the time the development work was initiated, this program, particularly the 1-kw part of it, benefited very much from the experience gained in various industrial and government laboratories with radar magnetrons within the 500-to-11,000-Mc range. The pulse operation required in radar equipment is particularly advantageous for oscillators of very high frequency. The conflicting conditions that must be satisfied in power tubes for very high frequencies have been pointed out in Chaps. 13 and 18. High voltage and small dimensions should be used to keep the time of transit of the electrons down, but high voltage leads to high power, because of the limited r-f impedances obtainable. Small dimen-

sions and high power are difficult to combine for two reasons: there are limits to the current density obtainable at the emitting cathode surface, and also to the power loss that can be dissipated in the tube without exceeding reasonable temperature limits. By the introduction of pulse operation, this dilemma was circumvented and its solution postponed until the time came to develop c-w magnetrons. Because in pulse operation the power is on only $\frac{1}{1000}$ of the time, say, the loss to be dissipated in the tube is reduced by the same factor of 1000. It was also discovered that oxide-coated cathodes could handle many times higher current densities under pulsed conditions than in continuous duty.

A considerable fund of knowledge and experience concerning pulsed magnetrons thus facilitated the design of c-w magnetrons. Two of the chief problems remaining to be solved were cathode design and heat dissipation. The demand for modulation introduced additional problems in the development of c-w tubes that were not encountered in pulsed tubes.

20-3. Main Features of a Magnetron Oscillator.—Before discussion of the various types of magnetrons and the principle of their operation, let us list the main component parts of all magnetron oscillators:

1. A cathode that emits electrons by thermionic or secondary emission, or a combination of both, of which the emitting part has the shape of either a cylinder or a filament, straight or helical
2. A cylindrical anode, coaxial with the cathode, usually divided into segments by slots parallel to the axis of the cylinder
3. A tank circuit connected to the anode segments and resonant at the desired frequency of oscillation, which incorporates means for coupling out power to a transmission line
4. A magnet to maintain a constant magnetic field parallel to the axis of the electrode system

20-4. Classification of Continuous-wave Magnetrons.—For historical reasons, two main categories of magnetron oscillator tubes are recognized: split-anode magnetrons and multicavity magnetrons. These names obviously refer to the number of anode segments, two or many (if the cavity aspect is disregarded for a moment), and may give the impression that there is a fundamental difference in the mechanism of oscillation in the two classes. It is now generally agreed that this is not so. There are other characteristics, however, that make this classification useful and significant. The split-anode tubes, for instance, in general have the glass envelope, filamentary cathode, and external tank circuit characteristic of the early tubes, while the multicavity tubes represent a radical break with the early tradition. There is a tendency, however, for new types of tubes to appear that combine various features from both

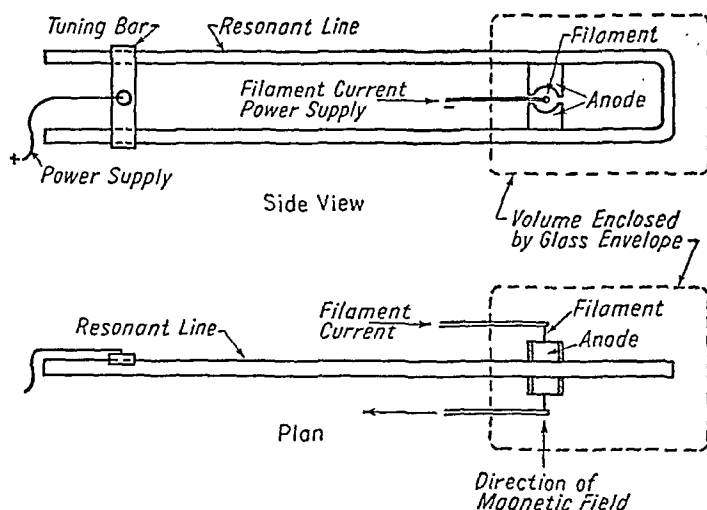


FIG. 20-1.—Schematic presentation of split-anode magnetron.

categories, as well as innovations that do not fall in either category. The neutrode and the finger magnetron, discussed in Sec. 23-2 and Sec. 22-12, respectively, are examples of this trend.

Figure 20-1 shows, somewhat schematically, a typical split-anode magnetron oscillator. A straight filament is placed along the axis of the cylinder formed by the inner surfaces of the two anode halves. The pole pieces of the magnet, as well as a substantial part of the parallel-line tank circuit, are external to the glass envelope. The tank circuit in this particular tube is "double-ended," *i.e.*, extended in both directions from the anode.

The modification known as a *neutrode* is sketched in the same fashion in Fig. 20-2. Here the anode is divided into three segments, but only two of them have any r-f potential with respect to the cathode, as the third segment is connected to a neutral point in the tank circuit.

Three types of multicavity magnetrons are shown in Fig. 20-3, differing only in the shape of the cavities. In all of them there is in the center a cylindrical oxide-coated cathode, provided with *end hats* or *end*

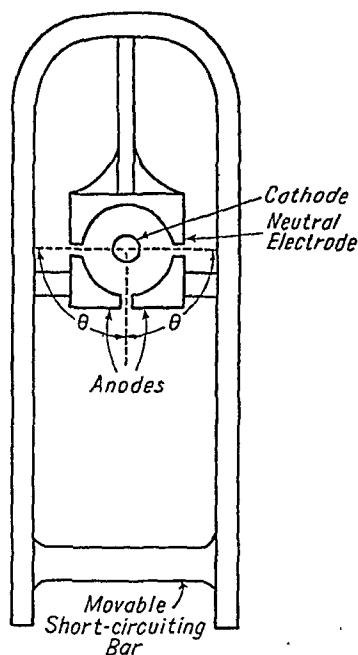


FIG. 20-2.—Magnetron with neutrode, neutrode.

shields to prevent electrons from leaving the *interaction space* between anode and cathode and leaking out in the axial direction. The pole tips of the magnet are built into the tube in order to keep the gap in the magnetic circuit to a minimum. The tank circuit of the oscillator is completely enclosed in the evacuated shell of the tube. In Fig. 20-3a it

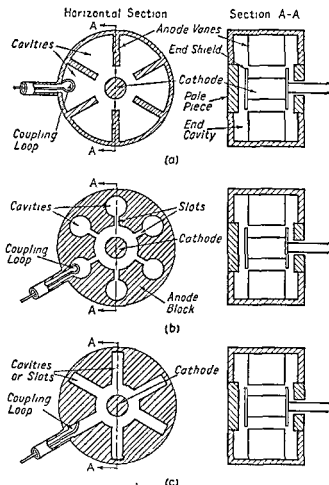


FIG. 20-3.—Multicavity magnetrons: (a) vane magnetron; (b) hole-and-slot magnetron; (c) slot magnetron.

is formed by the cavities between the six vanes and the end spaces into which these cavities open at each end in the axial direction. The coupling loop is the means for transfer of r-f energy from the cavity system to the output line. The tips of the vanes form the six anode segments. The pole pieces of the magnet are shown as integral parts of the shell of the tube. Figures 20-3b and c show differently shaped cavities, milled or drilled out of the *anode block*.

Figure 20-4 shows schematically a radical departure from the multicavity design: it is a single-cavity multisegment magnetron, referred to by different names as *finger magnetron*, *squirrel-cage magnetron*, *donutron*, etc. The interaction space and the electrodes are similar to those of a multicavity magnetron, but the anode segments form fingers, of which alternate ones are attached to opposite sides of a flat cylindrical cavity.

This chapter will cover the principles of operation chiefly of multicavity magnetrons. As already mentioned, the mechanism of oscillation of the split-anode magnetron is assumed to be the same. The difference between multicavity tubes and split-anode tubes is then mainly that the resonant system, *i.e.*, the tank circuit, of the latter is simpler. As

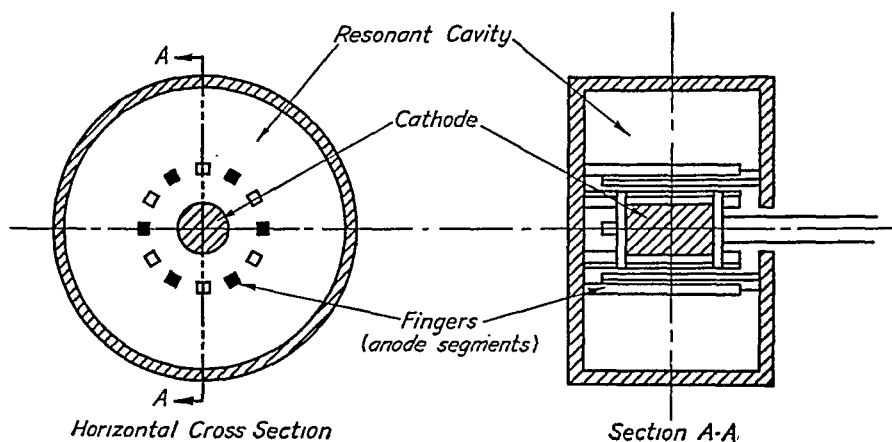


FIG. 20-4.—Donutron. For clarity all fingers connected to one side of the cavity have been shaded in the left section.

far as the fundamental operation is concerned, we can even consider the split-anode magnetron as a special case of a multicavity magnetron, so that by covering the principles of operation of the more general class of magnetrons we have automatically included the less inclusive category.

In order to describe how and why a magnetron operates, the following subjects will be taken up:

1. The configuration of the electromagnetic field in the region involving anode and cathode
2. The behavior of the rotating electron cloud that surrounds the cathode and serves as a means of converting part of the d-c input power into r-f output power
3. Cathode considerations, particularly with relation to scaling, emission-current density, back-heating, and factors affecting life
4. Output coupling for coaxial lines and waveguides with a discussion of conditions imposed by wide tuning range
5. Means for obtaining wide tuning range, consequent variations in shunt impedance, Q , and efficiency of the resonant system, and the problem of spurious resonances

The question of removal of heat from anode, cathode, and tuning components will be discussed in connection with the various tubes described in Chap. 22.

20-5. The Electromagnetic-field Configuration in a Multicavity Magnetron.—The problem is not so simple as in a triode or tetrode oscillator because in a magnetron the r-f potential is not constant over the whole anode surface, each segment having its own r-f potential. Because of the rather complicated shape of the interior of a multicavity magnetron, all parts of which are included in the resonant system or the tank circuit of the oscillator, a mathematical analysis of the possible electromagnetic-field configurations in this space is prohibitively laborious. However, the essential facts of the solutions and the underlying reasons therefor are not difficult to interpret in simple physical terms.

The procedure for a rigorous solution would be to integrate Maxwell's differential equations for the electromagnetic field over the whole space penetrated by the field. If infinite conductivity of copper is assumed, this integration will extend only over the vacuum in the cavity, and the boundary conditions everywhere at the metal surface will be that the tangential component of the electric field and the perpendicular component of the magnetic field must be zero. An infinite number of elementary solutions will be obtained, each representing a state of oscillation with a specific field pattern and a specific frequency. Because we have assumed no losses, these oscillations will have no attenuation as long as no coupling loop is inserted, and their amplitude and time phase will depend on the initial conditions only. The field configuration and frequency of each elementary solution is said to define a *resonance mode* of the cavity. When impressed oscillations are discussed the term *mode* will be used for a state of oscillation with the same field configuration, but not necessarily with the same frequency as one of these elementary solutions. When an a-c emf of a frequency different from any resonance frequency is exciting the cavity, the voltages, currents, and electromagnetic fields obtained will be much smaller than under resonance conditions, in complete analogy with a resonant circuit formed by lumped constants L , C , and R and excited off resonance. The electromagnetic field in the cavity can then be interpreted as a sum of different components each with the field configuration of one of the elementary modes. The above analogy can be extended to a number of resonant circuits with lumped constants, all connected together; the emf will produce a response in each circuit that depends on the reactance and resistance of the circuit at that particular frequency.

In describing some of these modes, it is convenient to deal separately with the annular space between anode and cathode, i.e., the *interaction space*, and the slots or cavities that separate the anode segments. The

width of the anode segments and the radial distance between anode and cathode is always small compared with the free-space wavelength of the oscillations considered here, so that the reactance of the small section of the interaction space under each segment will be unimportant compared with the reactance of the slot or cavity between two segments. It will thus not be far from the truth to assume that the cavities alone determine the frequencies of the resonance modes and that the field configuration of a certain mode in the interaction space is independent of the resonance frequency.

It is convenient to think of the electromagnetic-field pattern in the interaction space in terms of waves propagated around this annular

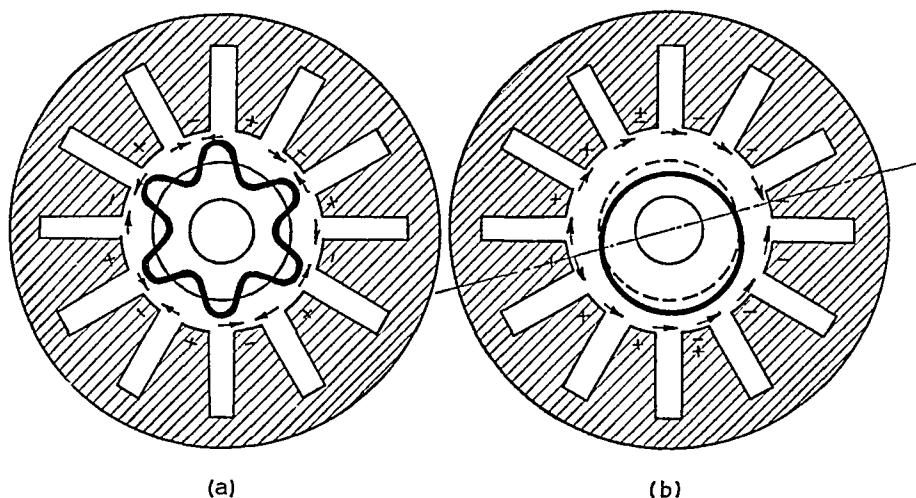


FIG. 20-5.—(a) Radial plot of the intensity of the principal component of the *tangential* electric-field strength in a 12-gap multicavity magnetron at resonance in the π -mode; (b) the corresponding plot for the mode, $n = 1$.

cavity. The velocity of propagation (more specifically, the phase velocity) is, according to our assumption regarding the dimensions of the interaction space, entirely determined by the phase delay occurring at each slot between the anode segments and by the frequency. In order to eliminate the frequency from our discussion until we proceed to the field in the slots and cavities, we can use one cycle as the unit of time. The velocity of propagation in these units becomes identical to the wavelength in the direction of propagation. (Note that this is not the free-space wavelength, but the interaction-space wavelength, *i.e.*, the actual length of one wave along the circumference of the interaction space.)

The field configurations in this annular space now have to satisfy the condition that the tangential electric field be zero along the surface of the anode segments and of the cathode. Our assumptions also demand that the potential be constant over each anode segment.

Figure 20-6 shows a plot of the tangential electric field and of the potential along a section of the cylindrical surface coinciding with the surface of the anode segments at a certain instant for the magnetron shown in Fig. 20-5. By something like a Fourier analysis of these rectangular and trapezoidal curves, we can find a number of sinusoidal components, all of which are complete electromagnetic waves consistent with Maxwell's equations and the sum of which satisfies the boundary

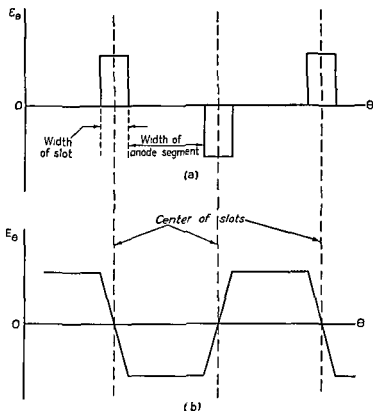


FIG. 20-6.—(a) Tangential component of electric field at anode, (b) r-f potential at anode.

conditions plotted in Fig. 20-6 and also gives zero tangential electric field at all points on the cathode. It might be well to emphasize that for each mode all these components have the same frequency; the Fourier expansion is in space, not in time. But the boundary conditions have to be satisfied at all instants during one complete cycle. The tangential component E_θ of the electric field at the anode radius is found to be

$$E_\theta = \sum_{m=0, \pm 1, \pm 2, \dots, \infty} A_m e^{j\left(\frac{2Y}{\lambda_p} + 2m\right)n\theta} \quad (20-1)$$

where Y is the distance between centers of adjacent slots, λ_p the wavelength measured along the anode circumference, θ the angular coordinate in a cylindrical coordinate system z, r, θ , coaxial with the anode and cathode, and n the mode number, *i.e.*, the number of wavelengths along the circumference of the anode. The mirror image of ε_θ in Eq. (20-1), which is obtained by replacing $+\theta$ by $-\theta$, is for reasons of symmetry also a satisfactory expansion.

Equation (20-1) represents one fundamental component $e^{j2Y/\lambda_p - n\theta}$ traveling in the positive θ direction [obtained for $m = 0$ in Eq. (20-1)] and an infinite number of components of shorter wavelengths traveling in both directions (obtained for $m \neq 0$). Thus to satisfy the boundary conditions requires, in addition to the forward wave, a reflected wave that is in general of smaller amplitude than the forward wave. If the wavelength is progressively reduced, this reflected wave increases in amplitude until it is equal to that of the forward wave when the wavelength is equal to twice the distance between slots. All the forward components will then be matched by reflected components of equal amplitude and wavelength, the component with $m = 0$ by the one with $m = -1$, the one with $m = 1$ by the one with $m = -2$, etc., and the result is a standing wave in the interaction space. We can compare the anode to the stator of a synchronous motor. As long as one wavelength is greater than twice the distance between slots, the anode acts as a polyphase winding, but when one wavelength just extends over two segments the anode acts as a single-phase winding. This latter case does not yet mean resonance in the slots or cavities; at such resonance the velocity and wavelength would be zero, if the losses were negligible.

So far the fact that waves propagated around the interaction space come back to the same point has been disregarded. All wavelengths can be propagated around the interaction space, but only waves of certain discrete wavelengths will come back in such a phase as to repeat themselves periodically. Only waves of those wavelengths will build up to appreciable amplitudes. These wavelengths then give the resonance modes. The longest resonance wavelength is equal to the circumference, and the corresponding mode is given the mode number $n = 1$. When two wavelengths equal the circumference, $n = 2$, and so on. The mode for which $n = N/2$, or half the number of anode segments, gives the standing wave condition mentioned above and is generally called the π -mode, because for this mode the phase difference between the field intensity of adjacent slots (or potential of adjacent segments) is exactly π radians. For all other modes than the π -mode the solution is degenerate, *i.e.*, two independent solutions exist, identical in all respects except the direction of rotation.

In Figs. 20-5a and b, the heavy lines represent the tangential-electric-field intensity plotted radially from a circular axis and show qualitatively how this field varies along the circle used as axis. The former figure concerns the π -mode ($n = 6$), and the latter the mode for which $n = 1$. In Fig. 20-7, the fundamental component of the tangential electric field is also plotted as a function of the radius with field intensity as ordinate

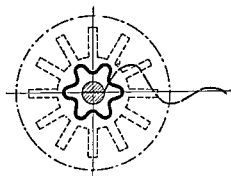


FIG. 20-7—Radial plot of the tangential field strength in a two-dimensional field pattern between the cathode cylinder and infinite space. The amplitude of the alternation in field strength varies radially according to the Bessel-function plot shown in heavy dashed lines

and radius as abscissa. This plot has been extended beyond the anode radius in order to demonstrate that this field component varies approximately as a Bessel function.

It is characteristic of all the modes that the amplitudes of the components of shorter interaction-space wavelength decrease much more quickly toward the cathode than the amplitude of the fundamental component, so that the latter becomes successively more predominant for decreasing radius.

It is therefore a fair approximation

to consider the fundamental component only, except for the immediate vicinity of the anode. We must remember, however, that the π -mode has two fundamental components of opposite direction of propagation.

So far the electromagnetic field in the slots or cavities between the anode segments has not been considered. It is rather easy to relate these parts of the resonant system to well-known circuit elements. In a hole-and-slot resonator (Fig. 20-3b) the slot is a capacitor and the hole a single-turn coil, both together forming a parallel-resonant circuit. A slot resonator (Fig. 20-3c) is a short-circuited transmission line, which also is a parallel-resonant circuit, if and when the length of the slot is equal to $\frac{1}{4}$ wavelength. A vane resonator (Fig. 20-3a) is something between these two extremes. In all these types of resonators, the magnetic-field lines are parallel to the anode axis and have their high density in the short-circuited or inductive end of the cavity. The electric-field lines are perpendicular to the walls of the slots and concentrated in the open or capacitive end of the cavity.

It would seem, then, that the interaction space, loaded at regular intervals with the resonant cavities, could be represented by an equivalent network like the one shown in Fig. 20-8a, where the capacitances AA' , BB' , etc., are the capacitances between the respective anode seg-

ments and the cathode, and the circuits $A'B'$, $B'C'$, etc., represent the impedances of the cavities between the anode segments.

In this picture, however, the fact that adjacent cavities are inductively coupled to each other has not been taken into account. This coupling exists because at the ends of the cavities in the axial direction the magnetic field lines bend over into the end spaces and return through neighboring cavities. In fact, the coupling is very tight, because only a very small fraction of the magnetic flux will return through the interaction space instead of through the other cavities. The equivalent network can easily be modified to take the mutual coupling into account by introducing the equivalent T of two coupled inductances (Figs. 20-8b and c). Because of the tight coupling the inductances L - M in the series arms can be neglected, and consequently Fig. 20-8a changes into 20-8d. There the shunt capacitances AA' , BB' have small susceptances compared with the inductances M , so that the equivalent network has the character of a high-pass filter. Consequently, it will propagate all frequencies above a certain cutoff frequency with negligible attenuation but with a certain phase shift for each section of the filter, i.e., for each slot between two anode segments. For very high frequencies this phase delay is negligible, but as the cutoff frequency is approached, the phase shift increases rapidly and reaches 180 deg at the cutoff frequency. The resonance frequency of the π -mode is thus equal to the cutoff frequency, and the resonance frequencies of the other modes, for which the phase difference between the electric field in consecutive slots is less than 180 deg, vary approximately as shown in Fig. 20-9. The π -mode has the lowest frequency, and the other modes have in consecutive order higher and higher frequencies. The frequency separation is smallest between the π -mode and the mode with next lower mode number. This *mode separation* is important for the stable operation of the magnetron. One per cent mode separation is considered poor, 3 per cent marginal, and 15 per cent excellent. It is

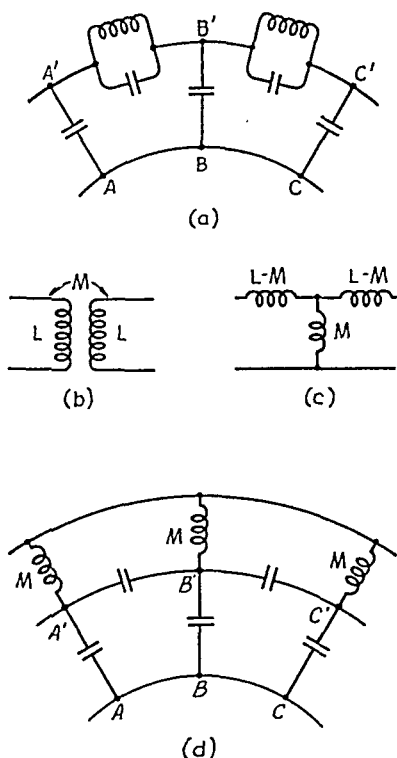


FIG. 20-8.—Representation of the anode system as a filter network.

usually necessary to increase the mode separation by artificial means, which will be described in Sec. 20-6.

It is important not to confuse the *component waves* and the *modes*. The presence of the component waves of any mode is required to satisfy the boundary conditions, and for each particular mode their relative amplitudes are always the same. Oscillations in the different modes, on the other hand, can be present in any amplitude and phase. In the case of forced oscillations, however, the modes may have a similar function to the component waves. If we introduce a loop in one of the slots and excite the system at a frequency different from any resonance frequency, the symmetry of

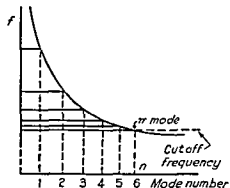
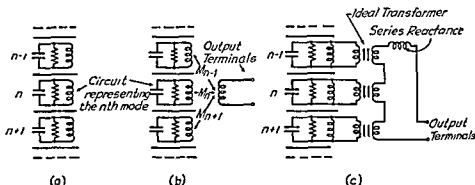


FIG. 20-9.—Frequency distribution of the modes of a 12-gap magnetron.

the system is disturbed and we need an expansion in terms of the different modes in order to satisfy the boundary conditions at the loop.

The amplitude and phase of oscillations in each mode caused by a sinusoidal emf vary with the frequency of this source approximately as



the oscillations in a simple resonant circuit. This is true even when the copper losses in the system are considered. As far as the variations with time and frequency are concerned, the system can therefore be represented by the *equivalent network* of Fig. 20-10a. As all the modes are orthogonal solutions,¹ there is no coupling between the circuits that repre-

¹ Orthogonal functions of *time* are familiar to every communications engineer; it is because sinusoidal oscillations of different frequencies are orthogonal that the currents and voltages of one frequency in a linear network are not affected if oscillations of other frequencies are superposed. Similarly, the different resonance modes of a cavity resonator are orthogonal functions of *space*; the currents, voltages, and fields

sent the different modes. A coupling loop, however, will couple to a large number of the modes (Fig. 20-10b). We can give the equivalent network a more convenient form by replacing the coupling loop with an ideal transformer and a certain series reactance (Fig. 20-10c).

20-6. Strapping.—As shown in Fig. 20-10a, there is one set of *equivalent circuit constants*, inductance, capacitance, and conductance, for each mode, as measured from two adjacent anode segments, for instance. These concepts are very helpful in the discussion of methods for increasing the mode separation. If we add a capacitance to the equivalent circuit, as shown in Fig. 20-11a, it will in general lower the resonance frequency of the mode; the higher the impedance between the points where the capacitance is introduced, the larger will be the change in frequency obtained. On the other hand,

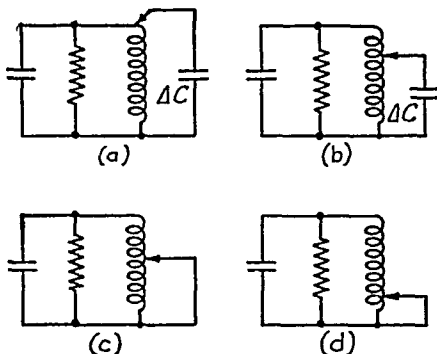


FIG. 20-11.—Illustrations of methods for increasing mode separation: (a) maximum frequency change; (b) less than maximum frequency change; (c) large frequency change; (d) small frequency change.

if we by-pass part of the equivalent inductance by means of a low-inductance lead, as shown in Fig. 20-11c, the frequency will increase. This increase will be greater the larger the initial impedance between the points where the by-pass is applied.

The most common method of increasing the mode separation is *double-ring strapping*, illustrated in Fig. 20-12. Two rings are attached to the anode segments or to the vanes close to their tips in such a way that alternate segments are connected to the outside ring and to the inside ring. In the π -mode the voltage on all the segments connected by one ring is in opposite phase to the voltage on the other ring and the other set of segments. The capacitance added by the rings will thus be most effective in lowering the resonance frequency of the π -mode. Other modes

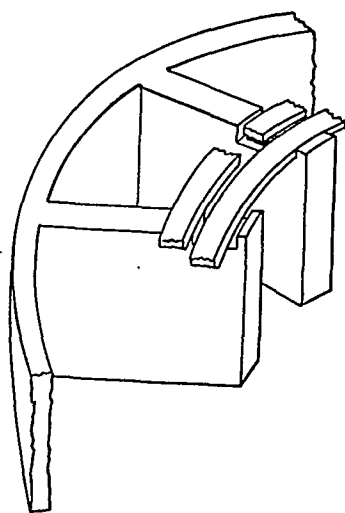


FIG. 20-12.—Double-ring strapping.

of one mode are not affected by the presence of currents, voltages, and fields of other modes, even if the impressed emfs in all modes have the same frequency.

for which the phase difference between adjacent segments is smaller will be less affected by this capacitance. Since in the π -mode all points connected by the same ring are at the same potential, no by-passing effect is produced, and the equivalent inductance remains unchanged. For the other modes all the points joined by the rings have different potentials. Heavy short-circuit currents will flow, and the equivalent inductance of these modes is reduced. Both the capacitance and the inductance are thus affected in such a way as to increase the frequency separation between the π -mode and the other modes. The capacitance between the strap rings plus the capacitance of each ring to the opposite set of anode segments is called the *strap capacitance*. The ratio of the strap capacitance to the equivalent capacitance of the π -mode is conventionally used as a measure of the strapping. The shorter and heavier the strapping segments are, the more tightly strapped the magnetron is said to be.

20-7. Circuit Constants of the Equivalent Network.—The resonant system of a multicavity magnetron, or more specifically, the π -mode of the system, is subject to very similar considerations to those mentioned in connection with the tank circuit of triode and tetrode oscillators as far as shunt impedance, Q , circuit efficiency, etc., are concerned (Secs. 18-6, 18-7, 18-10).

For the purpose of this discussion we simplify the equivalent circuit to the form shown in Fig. 20-13, which will be the standard form of equivalent circuit for a magnetron resonator in this and the next chapter.

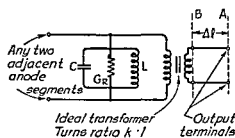


FIG. 20-13—Equivalent circuit for operation in the π -mode.

This circuit is based on the following assumptions: the Q of the π -mode is high, and we are interested only in a small frequency range at resonance, a few per cent at the most. The mode separation is satisfactory, so that in this frequency range the other modes contribute only very small reactances.

The output-coupling network does not contain any resonant elements, and its losses are negligible. If the tube is tunable, in general all constants of the equivalent circuit will vary with the setting of the tuner, but for each setting the above assumptions should be correct.

Let us now define some of the important parameters of the tank circuit in terms of the equivalent inductance L , capacitance C , conductance G_r , the turns ratio k of the equivalent ideal transformer of the output coupling, the equivalent length of line Δl of the output coupling, the characteristic impedance Z_0 of the output transmission line, and the angular frequency ω .

The unloaded Q is

$$Q_0 = \frac{1}{\omega L G_r} \quad (20-2)$$

The loaded Q , which is conventionally defined for matched load on the output line (load impedance $Z_L = Z_0$), is

$$Q_L = \frac{1}{\omega L \left(G_r + \frac{k^2}{Z_0} \right)} \quad (20-3)$$

The *external* Q , i.e., the Q we would get with the same load $Z_L = Z_0$ but no losses in the resonator ($G_r = 0$), we obtain from the relation

$$\frac{1}{Q_E} = \frac{1}{Q_L} - \frac{1}{Q_0} \quad (20-4)$$

The circuit efficiency

$$\eta_c = \frac{Q_L}{Q_E} = 1 - \frac{Q_L}{Q_0} \quad (20-5)$$

The slot admittance measured between adjacent anode segments for matched load on the output line is

$$Y_i = \frac{k^2}{Z_0} + \frac{1}{Q_0} \sqrt{\frac{C}{L}} = \frac{k^2}{\eta_c Z_0} \quad (20-6)$$

since $(1/Q_0) \sqrt{C/L} = G_r$. The reciprocal of this admittance corresponds to the shunt impedance discussed in previous sections on triode and tetrode oscillators (see Sec. 18-6).

The equivalent-circuit constants and the parameters derived from them can be measured relatively easily. Impedance measurements from the output terminals, usually referred to as *cold tests*, yield the Q 's and η_c explicitly and Y_i , L , C , and G_r in terms of k . In a demountable-tube model, k can be measured by increasing the equivalent capacitance by a known amount and again determining the impedance at the output terminals. The cold tests will be discussed further in Secs. 24-29 to 24-35.

It is often convenient to refer the equivalent-circuit constants to the output terminals and to *normalize* them, i.e., measure them in terms of the characteristic impedance Z_0 of the output line. Using the subscript

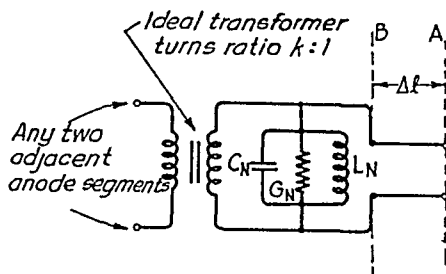


FIG. 20-14.—Equivalent circuit normalized with respect to the output line.

N to indicate this transformation of the constants, we obtain

$$L_N = \frac{L}{k^2 Z_0} = \frac{1}{Q_{E\omega}} \quad (20-7)$$

$$C_N = k^2 Z_0 C \quad (20-8)$$

$$G_N = k^2 Z_0 G, \quad (20-9)$$

We can redraw the equivalent circuit in terms of these normalized constants, as shown in Fig. 20-14.

It is apparent from Eq. (20-7) that the external Q is equal to the reciprocal of the normalized inductive reactance of the tank circuit as measured from the output terminals, which makes it a very useful quantity.

20-8. The Electron Mechanism of Oscillation.—The behavior of the space charge in a magnetron is extremely complicated, and no quantitatively satisfactory theory has yet been worked out. Qualitative theoretical considerations and extensive experimental information, however, enable us to describe what happens in the interaction space of an oscillating magnetron.

It is instructive to consider the motion of the electrons in a magnetron when no oscillations take place. All the anode segments have the same potential, and all conditions will be approximately the same as if the anode had a continuous cylindrical surface. Under the influence of a constant axial magnetic field \mathcal{H} or flux density \mathfrak{B} and a constant anode voltage E_b , an electron will describe a circle with radius r , concentric with the cathode if its velocity v is such that

$$\mathcal{E}e + \frac{mv^2}{r} = \mathfrak{B}ev = \mu_0 \mathcal{H}cv \quad (20-10)$$

where \mathcal{E} is the field intensity due to E_b , μ_0 is the permeability of free space, and e and m are the charge and mass of the electron, respectively. This relation simply states that the sum of the force exerted by the electric field on the electron and the centrifugal force due to the curvature of the electron orbit equals the force caused by the motion of the electron through the magnetic field, all in mks units. Provided the plate voltage is smaller than the *critical voltage* or *cutoff voltage*

$$E_c = 2\pi^2 \frac{e}{m} \mathcal{H}^2 r_a^2 \left[1 - \left(\frac{r_c}{r_a} \right)^2 \right]^2 10^{-14} \quad \text{volts} \quad (20-11)$$

no electrons will reach the anode. Here r_a and r_c are the radii of anode and cathode, respectively, in meters. It is in this case perfectly possible that all the electrons describe circular orbits and form a uniform space charge revolving around the cathode. The velocity of the electrons would grow from zero at the cathode to a maximum value at the edge of

the electron cloud, in such a way that Eq. (20-10) is satisfied at every point. The potential at the edge of the cloud would be obtained from Eq. (20-11) by replacing the anode radius with the radius of the cloud.

Other electron orbits can also build up consistent solutions for the space charge of a nonoscillatory magnetron, but it has been shown that the average space-charge density and average angular velocity of the electrons are always approximately the same as for the simpler solution described above. We are, therefore, justified in using this simple picture as a starting point for a discussion of the electronic behavior of the magnetron.

In order to explain why and how a magnetron oscillates, we should show two things. (1) The cloud of electrons can deliver power to a finite, steady-state r-f voltage in the tank circuit, and thus support the oscillations in this circuit. This power would, of course, have to be taken from the d-c power supply by electrons moving from cathode to anode and establishing a direct current. (2) This condition of power transfer can be maintained continuously for a range of r-f voltage from infinitesimal amplitude up to normal amplitudes, so that the oscillations are self-starting.

If we excite the anode cavity system in the π -mode, the electrons in the cloud will be subject to an r-f electric field in the direction of their motion, according to Fig. 20-7, and consequently be accelerated at some spots and retarded at other spots. Of the component waves associated with the π -mode the fundamental component that moves in the same direction as the electrons is the most important. If the anode voltage and the magnetic field are so adjusted that the electrons at the edge of the cloud are synchronous with this component, these electrons will remain under the influence of the r-f field for considerable time. In a 12-cavity magnetron shown in Fig. 20-7, there will be six regions of acceleration and six regions of retardation. The effect on the angular velocity will be a bunching of the electrons at the rim of the cloud. Where the velocity has been reduced, the magnetic force $\mathcal{B}ev$ is reduced so that the force of the electric field $\mathcal{E}e$ will predominate and pull the electrons toward the anode. The electrons will regain their velocity from the d-c field but at the same time move into a region of stronger r-f field. They will thus continue proceeding toward the anode, extracting energy from the d-c field and losing energy to the retarding r-f field. On the other hand, where the electrons are accelerated, the magnetic force predominates and pulls the electrons toward the cathode, out of reach of the r-f field. The space-charge cloud will consequently assume a quite different shape. The "hub" of space charge around the cathode will be reduced in diameter, while the groups of decelerated electrons reach out as the spokes of a wheel toward the anode. The "spokes"

will continually revolve against the force exerted on them by the r-f field, thus feeding power into this field and maintaining the oscillations in the tank circuit. On the other hand, very little r-f power is required to keep electrons out of the regions where acceleration would take place. It is probably this selective mechanism that is ultimately responsible for the fact that surprisingly high electronic efficiencies, sometimes over 80 per cent, are experimentally obtainable in magnetrons.

As far as the question of starting the oscillations is concerned, there is nothing in the exchange of energy as described above that demands a large r-f voltage on the anode. This mechanism can consequently start to function at infinitesimal r-f voltage and gradually build up the oscillations in the tank circuit.

It is typical of magnetron-oscillator behavior that, at a given magnetic field, operation at a given frequency and mode cannot occur until the anode voltage has reached a more or less definite critical value, sometimes called the *threshold voltage*. This critical voltage is found experimentally to be approximately proportional to the magnetic-field strength. If the threshold voltage is plotted for the various modes of a magnetron vs. the magnetic field, a diagram of the character shown in Fig. 20-15 is obtained. The cutoff parabola according to Eq. (20-11) is shown at the left, and

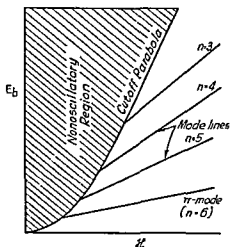


FIG. 20-15.—Mode-line diagram.

the lines of the threshold voltage for the different modes, often referred to as *mode lines*, start at this curve. For a given geometry of the interaction space the position of each line depends on the angular velocity of the fundamental component wave of the corresponding mode. An increase in the angular velocity will move the mode line toward higher anode voltages. As the angular velocity is proportional to f/n , the mode line will move upward with increased resonance frequency and downward with increased mode number n . The distance between the mode lines

will consequently decrease with increased number of segments, because both the difference between the highest mode numbers ($= 1/n$) and the mode separation in frequency will decrease with increased n . It should be noted that tuning the magnetron to a different frequency will move the mode line, upward if the frequency is increased, downward reduced.

The previous section stated that the fundamental component of each mode is predominant over most of the interaction space, and it has consequently been assumed that the space charge synchronizes with the fundamental component of each mode. This is not necessarily always true; successful operation of 10,000-Mc tubes has been reported in which the space charge synchronizes with the lowest order component of opposite direction of rotation to the fundamental component of a mode of smaller mode number than the π -mode. The fundamental component of the mode thus in this case rotates in the opposite direction to the electrons. The efficiency obtainable in this kind of operation is lower than when the fundamental component is used, but angular velocity of the electrons and the anode voltage required may be even lower (or the dimensions larger for the same voltages) than those of the π -mode. The same resonance mode can in this way produce more than one line in the mode-line diagram, and one of these may be even lower in the diagram than the line representing the π -mode.

No discussion of the modern magnetron would be complete without a few words about *magnetron scaling*. This powerful method combined with a wealth of experimental data on successful magnetron designs has proved a very useful substitute for a quantitatively satisfactory theory, for design of new magnetrons as well as for critical evaluation of existing tubes. The basis of scaling is the fact that it is possible to imitate a certain model design with different dimensions in such a way that some important characteristics—the efficiency of a magnetron, for instance—are carried over unchanged to the new design. If a number of variables is involved, which are related by nonlinear equations, the scaling process is, of course, not quite so simple as to reproduce a drawing or a map to a different scale. Still, it is possible to analyze these equations and deduce the scaling rules without the necessity of finding a solution to the equations.

For magnetrons several kinds of scaling are possible: wavelength scaling, voltage scaling, and current scaling. The purpose of wavelength scaling is obvious; the choice of voltage and current, on the other hand, is subject to other factors than power output, such as resonator design, circuit efficiency, cathode emission, and practical considerations as to power supply.

Slater has condensed the rules for magnetron scaling into expressions for three quantities E_0 , \mathcal{C}_0 , I_0 by means of which the anode voltage E_b , the anode current I_b , and the magnetic field \mathcal{C} can be transformed to dimensionless variables E_b/E_0 , $\mathcal{C}/\mathcal{C}_0$, I_b/I_0 . At the same operating point in these variables all magnetrons have the same efficiency, provided that the shunt impedance of the tank circuit divided by E_0/I_0 is the

same and that the geometry is similar in all respects not taken into account in the calculation of E_0 , \mathcal{H}_0 , I_0 .

The expressions for these quantities are

$$E_0 = 253,000 \left(\frac{2\pi r_a}{n\lambda} \right)^2 \quad \text{volts} \quad (20-12)$$

$$\mathcal{H}_0 = \frac{16,900}{n\lambda \left[1 - \left(\frac{r_c}{r_a} \right)^2 \right]} \quad \text{amp turns per m} \quad (20-13)$$

$$I_0 = 8440 \frac{a_1}{\left[1 - \left(\frac{r_c}{r_a} \right)^2 \right]^{\frac{1}{2}} \left[\frac{r_a}{r_c} + 1 \right]} \left(\frac{2\pi r_a}{n\lambda} \right)^3 \frac{l}{r_a} \quad \text{amp} \quad (20-14)$$

where n = the mode number, i.e., the interaction-space circumference, divided by the interaction-space wavelength

r_a = anode radius, meters

r_c = cathode radius, meters

l = the axial length of the electrode system, meters

a_1 = the geometry constant in Langmuir's equation for the space-charge current in a cylindrical diode (≈ 0.95).

The first relation shows what to do if a magnetron is to be scaled to another voltage. The simplest way is to change the anode radius as the square root of the ratio of the new voltage to the previous voltage. Because for the π -mode $n = N/2$, i.e., half the number of cavities, voltage scaling can also be accomplished by changing N . The rule is then evidently that r_a/N should be changed in the ratio of the square roots of the voltages. If the latter procedure is followed, both magnetic field and current will be affected, if the former, only the current.

In current scaling, Eq. (20-14) states that a different current can be obtained at constant efficiency by varying the ratios r_a/r_c or l/r_a without any corresponding change in voltage or magnetic field. There are, however, certain limitations in the choice of r_c and l , such as current density at the emitting surface and cooling of the electrodes, that have to be taken into account.

In Eqs. (20-12) and (20-14), the product of the wavelength by the mode number and the linear dimensions appear only in ratio; wavelength scaling consequently requires that the linear dimensions be varied in direct proportion to this product. Both voltage and current will remain unaffected, but according to Eq. (20-13) the magnetic field required for equivalent operation will change.

It is interesting to consider the question of tunable magnetrons in the light of these rules for wavelength scaling. The dimensions of interaction space are not changed when the tube is tuned; consequently it

should be necessary to vary E_b , \mathcal{C} , and I_b in direct proportion to E_0 , \mathcal{C}_0 , and I_0 in order to maintain the dimensionless variables and the efficiency constant. Equations (20-12) to (20-14) show that for a tube with a tuning range of 1.5 to 1 the voltage would thus have to be varied over the tuning range by a factor of 2.25, the magnetic field by 1.5, the current by 3.38, and the shunt impedance by 1.5, with the result that the output power would vary by a factor of 7.6. Fortunately the region of fair-efficiency operation, expressed in Slater's dimensionless variables, is rather wide, so that it is by no means necessary to maintain the same operating point when the wavelength is varied. However, the discussion of some of the wide-range tunable tubes later on will show that the violation of the scaling rules in tuning the tubes has some important consequences.

20-9. Cathodes for Continuous-wave Magnetrons.—The operating conditions of the cathode in a magnetron are rather different from those of a cathode in a triode or a tetrode. The most characteristic feature is the electron bombardment to which the cathode is exposed. A small fraction of the electrons in the space-charge cloud will always be accelerated by the r-f field, return to the cathode, and land there with varying amounts of kinetic energy.

This bombardment has two important consequences: it raises the temperature of the cathode, and it produces secondary emission. The former phenomenon, which is called *back-heating*, is a very important factor as far as the life of the cathode is concerned. The secondary emission, on the other hand, increases the emission capability of the cathode and is thus a favorable factor. One of the magnetrons described in Chap. 22 is an example of a pure secondary-emission cathode with only a small thermionic filament to initiate the oscillations.

The design and performance of magnetron cathodes is intimately connected with the scaling of the tube. The modern trend has been to design tubes for successively shorter wavelengths. At constant voltage and current, the scaling rules, as explained in a previous section, demand that a reduction in wavelength be accompanied by a reduction of the linear dimensions or an increase of the number of cavities, or both. The increasing difficulty of obtaining adequate mode separation prohibits an increase of the number of cavities beyond certain limits, so that the higher the frequency, the smaller the cathode dimensions will have to be made. The small size of the cathodes leads to high current density and strong back-heating. To design cathodes that meet these severe requirements is one of the hardest problems concerning 10,000-Mc c-w magnetrons, and even in 3000-Mc magnetrons of 1 kw and above, the cathode is the main factor that determines the length of the life of the tube.

Cathodes of many different materials and shapes have been used in

the various c-w magnetrons that have been developed up to the present time.

The regular oxide-coated nickel cylinder is an excellent emitter but rather vulnerable to sparking and arcing. Nickel-mesh and nickelized-oxide cathodes stand up better under such conditions but have somewhat lower emission. Serious doubts have been expressed as to the ability of oxide cathodes to perform creditably in c-w magnetrons at higher voltages, 3000 to 5000 volts. There is recent evidence, however, that a modern nickel-mesh cathode at reasonable current density will have a satisfactory life at voltages of this order.

Experiments with tungsten filaments embedded in thoria gave promising results, but satisfactory life and stability do not yet appear to have been obtained. A 1000-watt 3000-Mc tube using a carburized thoriated helical filament has reached the final stages of development and will be described in Chap. 22.

Plain tungsten filaments are still used to a large extent, either straight or helical for the split-anode magnetrons, helical for a few multicavity tubes. A radical departure from conventional design is a tantalum sleeve acting as cathode and carrying heating current longitudinally. A tungsten rod along the axis of the sleeve serves as return path for the heating current. This cathode was developed for a 10,000-Mc tube where all ordinary cathodes were found inadequate owing to the high current density and consequent severe back-heating.

In many tubes where back-heating may seriously limit the life of the tube, some steps must be taken to protect the cathode against excessive temperatures. One way of doing this is to reduce the heating current when back-heating occurs, so as to maintain the correct temperature. For filamentary cathodes the resistance is a sensitive indication of the temperature, and automatic regulators that vary the heating power in such a way as to keep the resistance of the filament constant have been built. One type of regulator for this purpose will be described in connection with the 6J21 tube in Chap. 22.

Obviously the range of these automatic devices is limited by the heating power. If the back-heating power exceeds the normal heating power, all the regulator can do is to disconnect the heating power entirely, and possibly interrupt the operation of the oscillator by means of suitable interlocking relays. The cathode must be designed so that the heating power in normal operation is larger than the back-heating power. Thus here the unusual condition prevails that the high power consumption of incandescent tungsten and tantalum cathodes is an advantage instead of a disadvantage, at least as far as cathodes with high current density and back-heating are concerned. The fact that a large fraction of the heat transfer is by radiation and consequently varies very rapidly with tem-

perature is an additional advantage of high-temperature cathodes. The design also aims to achieve the best possible conduction cooling of the cathode. Water cooling has so far been limited to secondary-emission cathodes; an example is the cathode of the 10-kw tube described in Sec. 22-2.

Among the factors affecting the life of the cathode, the temperature is the most important; rigid control of the temperature is thus imperative in all tubes where back-heating is appreciable. A thoriated-tungsten filament is especially susceptible to temporary or permanent stripping of its active layer if operated at too high current density, so that a voltage gradient appears at the cathode surface. There is also some experimental evidence that modulation under certain conditions can adversely affect the life of a magnetron cathode. A modulation voltage is applied in series with the d-c voltage, and, because of the nonlinear character of the tube characteristic, peak anode current of more than three times the average currents may be obtained. It is an empirical fact that for each mode and magnetic-field intensity there is a maximum value of anode current. If the voltage is further increased after this current value has been reached, the oscillations cease; oscillations may, and usually do, at the same time start in a different mode at lower current. For this reason this current limit is often referred to as *mode-jump current*. Abnormally short life has been observed when modulation has caused shift of mode at every modulation peak, but the relation between cause and effect has not been clearly established. There is no doubt, however, about the fact that when the emission of the cathode has deteriorated beyond a certain point, the mode-jump current starts to drop off. In applications requiring modulation, this is actually the criterion for the end of the useful life of a c-w magnetron.

20-10. Output Coupling.—The purpose of the output coupling is to transmit r-f power from the tank circuit to the transmission line leading to the load. In a previous section (Sec. 20-7) the tank circuit has been discussed and the fact mentioned that the output coupling determines the loaded Q and the circuit efficiency of the resonant system. In multicavity magnetrons, where the output coupling, like the tank circuit, is built into the tube and cannot be adjusted from the outside, the design of the coupling element is of considerable importance and must be carried out with full appreciation of the operating conditions and required performance in the practical application of the tube. For the equipment designer, it is equally important to realize the limitations imposed by the fixed coupling. This subject will be discussed in Sec. 21-2 and 21-6 in connection with the operating characteristics of magnetrons.

In this section our purpose is to describe some methods for coupling out power in magnetrons and to discuss briefly how the power transfer

varies over the tuning range of a broadly tunable tube. Figure 20-16 illustrates a number of coupling schemes for magnetrons. They can be classified according to the nature of the transmission line or of the coupling element itself. In the former case we distinguish between *coaxial-line output* and *waveguide output*, in the latter between *loops* and *apertures*.

Figure 20-16 shows two arrangements of a coupling loop for coaxial-line output. The loop is placed where the r-f magnetic field is intense,

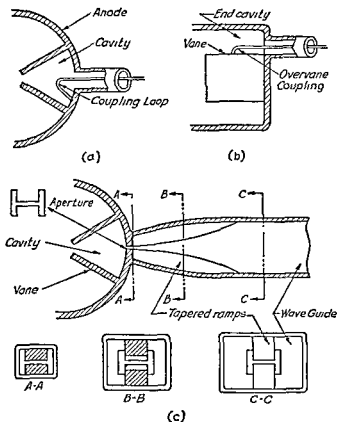


FIG. 20-16 — Methods for coupling out power.

and it is so oriented that the plane of the loop intercepts the flux lines perpendicularly. In Fig. 20-16a the loop is mounted in the back of one of the cavities in the anode block; in *b* its position is in the end cavity in the extended plane of one of the anode vanes. For a given cavity and a given transmission line the power transfer is affected by two characteristics of the loop, its area and its uncoupled, or leakage, inductance. The most uniform power transfer over the tuning range is obtained if the

latter is as small as possible. For this reason the loop is usually formed from a broad strip of copper, and the wall of the cavity or the surface of a vane is used as part of the loop.

Figure 20-16c shows an example of an aperture coupling for waveguide output (see also Fig. 22-4). The aperture between the resonant system and the end of the waveguide resembles an H; the vertical lines of the H correspond to broad openings, one into each end cavity, and the horizontal line to a narrow slot through the back of one of the cavities in the anode block. If an ordinary rectangular waveguide were used all the way up to this window, the power transfer would vary very rapidly with the frequency. This difficulty is avoided by making the first part of the waveguide of the same shape as the aperture and connecting this waveguide to the standard waveguide by means of a tapered section in which shape and dimensions gradually change to match the H section at one end to the rectangular section at the other end. The characteristic impedance of the H is chiefly determined by the width of the horizontal slot, while the purpose of the vertical parts is to maintain a sufficiently high cutoff wavelength. The shape and spacing of the "ramps" that start at the horizontal slot thus largely determine the impedance introduced into the cavity by the output coupling and the variation of this impedance with frequency. As the width of the slot is very small, a high precision is necessary. The manufacturing tolerances of the "horn" that forms the remaining boundaries of the tapered section are much less stringent.

A similar output system for coaxial-line output has been developed experimentally but has not so far found use in tube production.

20-11. Achievement of Wide Tuning Range.—In this section also, only magnetrons with internal resonators, primarily of the multicavity type, will be considered. Chapters 22 and 23 will give examples and discussions of other tunable oscillators.

The frequency of a magnetron oscillator can be varied (1) by modifying the geometry of all the cavities in the anode block, (2) by "pulling" the frequency by means of a tunable cavity that is coupled to the tank circuit, and (3) by changing the effective capacitance in the tank circuit or in a resonator coupled to it by means of a cloud or beam of electrons. The second and third methods, as well as others based on thermal expansion, etc., have not yet given sufficient tuning ranges in c-w magnetrons to be of any interest here and will therefore not be discussed further.

Modification of the geometry of the individual cavities can affect the equivalent capacitance, the equivalent inductance, or both. In the most common arrangement this is accomplished by mechanical movement of metal rings inside the vacuum chamber. The lower power magnetrons (50 to 200 watts) employ only rings that affect the capacitance; these are called *C-rings*. In the 1-kw magnetrons both capaci-

tance-modifying and inductance-modifying rings (*C*-rings and *L*-rings) are used.

The simplest *C*-ring arrangement consists of a thin flat ring of rectangular cross section placed close to and coaxial with the strap rings on one side of the anode block in a double-ring-strapped magnetron. This *C*-ring is almost in contact with the strap rings in the low-frequency position, but is a substantial distance from them in the high-frequency position.

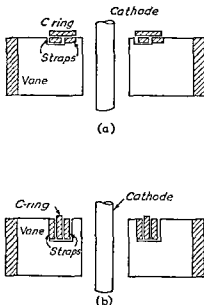


FIG. 20-17.—*C*-ring tuning.

C-ring is close to the straps produces a much larger effect than a movement of the same amount when the *C*-ring is moved away a distance comparable with the width of a strap.

A distinct improvement in tuning range is obtained if both straps and *C*-rings are cylindrical instead of flat. The straps are mounted in a slot in the vanes, and the *C*-ring moves in between them, as illustrated in Fig. 20-17*b*. Because of the shape of the *C*-ring, this arrangement has been called "cookie-cutter" tuning. Here the variable capacitance is a greater proportion of the total than with the flat *C*-ring, and the frequency is more nearly a linear function of the tuner position. These advantages are gained at the expense of reduced rate of heat removal from the vane tips beyond the rings. Tunability obtained in this way has ranged up to a ratio of 1.13:1. The possible tuning range is limited by the voltage between the straps and the smallest clearance that can be realized with reasonable manufacturing tolerances. In low-voltage tubes the last-mentioned limitation is the most important.

Capacitance tuning has the drawback that it is accompanied by a change of the loaded *Q*. For this reason it is desirable to vary also the

inductance. The simplest approach to this is to provide a flat ring overlying the back ends of the cavities. When this L -ring is close to the edges of the vanes, it seriously impedes the passage of r-f magnetic flux from one cavity to the next, and so reduces the inductance. If a C -ring and an L -ring are used in combination and fastened to a common movable support, their effect on the frequency of oscillation will add. When the C -ring moves close to the straps on one side of the anode block, the L -ring moves away from the vanes on the opposite side of the block; thus both L and C increase. This arrangement increases the obtainable tuning range, tends to keep the loaded Q more nearly constant by changing L and C in nearly like proportions, and improves the linearity of the tuning curve, at least as compared with the design illustrated in Fig. 20-17a. An early design with flat L - and C -rings is shown in Fig. 20-18. An important improvement has been made in the tuner employed in the 6J21 magnetrons (see Fig. 22-4). The straps at the C -ring have been omitted; both L - and C -rings are of heavy construction and have deep slots to clear the vanes, so that the rings can penetrate deep into the cavities without touching the vanes. This design has the advantage of the "cookie-cutter" rings as far as linearity of tuning is concerned, but not their disadvantage, as the flow of heat in the vanes is unobstructed. A tuning range of 1.5 to 1 has been consistently obtained with this tuner, and dependable operation has been achieved at all frequencies within this range.

A method of tuning employing what has been called the *crown-of-thorns* principle has been developed for radar magnetrons in the 10,000-Mc frequency range and has also been used in the 10,000-Mc work on the c-w tubes of the 1-kw class.

The *crown of thorns* itself is a set of rods, one to each cavity, which lie parallel to the cathode axis, have such cross sections that each more or less completely fills one cavity, and are moved together parallel to themselves in or out of the cavities to accomplish the tuning. Figure 20-19 illustrates the principle of the crown-of-thorns tuner. It operates similarly to the slotted L -ring in the 6J21 tube; the narrow gap between each thorn and the walls of the cavity introduces a high reluctance in the r-f magnetic circuit and reduces the flux and consequently the inductance. A c-w magnetron has been built with this type of tuner, which gave a 1.5-to-1 frequency range with its center at approximately 9000 Mc. The curve of frequency vs. tuner position had a desirably close approach to linearity.

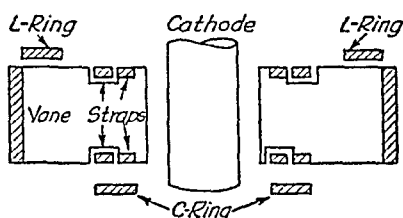


FIG. 20-18.—Tuning by means of L - and C -rings.

The crown-of-thorns construction has some other advantages, particularly as to the likelihood of spurious resonances, which cause many difficulties in tubes with a rather complicated structure to support, move, and guide the rings. In such a structure all dimensions cannot be kept small compared with the wavelength, nor all structural parts adequately by-passed to ground. The method of attack in such tubes is therefore not to eliminate the spurious resonances, but to render them harmless. This can be done by reducing the coupling between the main resonant system and the spurious systems and by making sure that the tuning curves of the spurious resonances do not intersect the tuning curve of the π -mode, because such intersections are liable to cause points of low

efficiency in the frequency range and discontinuities in the tuning curve (see Sec. 22-3). A description of the structure of the crown-of-thorns tuner will reveal the advantages of this system with respect to structural simplicity. The rods are supported on a mounting block that slides in, and is completely surrounded by, a metal casing; the thorns protrude through the end of this casing, which forms the wall of the end cavity. As the casing con-

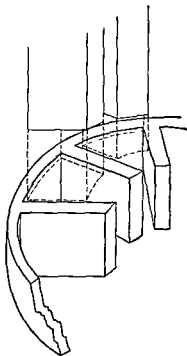


FIG. 20-19—Crown-of-thorns tuning scheme.

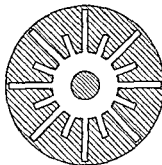


FIG. 20-20—Rising-sun anode

sequently does not move, the distance from the anode block to all other metallic parts remains constant; only the length of the thorns changes. The clearances between the thorns and the casing are very small and form a satisfactory by-pass. The shape and dimensions of all space reached by the r-f field are relatively simple and small, and the risk of trouble from spurious resonant systems is therefore minimized. The small clearances required to maintain this advantage, of course, introduce difficulties of construction and manufacture.

In the early stages of the design of the 10,000-Mc magnetron referred

to above, consideration was given to the possibility of using the "rising-sun" type of anode, illustrated in Fig. 20-20. This type of resonant system can provide adequate mode separation without strapping and has been successfully used in radar magnetrons. When the resonance frequencies of the modes of a resonant system of this kind are measured and plotted as in Fig. 20-9, it is found that the curve has an appreciable discontinuity at one point, *i.e.*, that the frequency difference between two particular consecutive modes is appreciably larger than between consecutive modes elsewhere in the diagram. For a certain relation between the lengths of alternate slots, this break in the curve can be made to fall between the π -mode and its neighbor, so that stable operation in the π -mode is secured. However, in order to obtain adequate tuning range and maintain satisfactory mode separation at all frequencies, one would have to tune both the long and the short cavities, a process that has so far been considered too complicated to warrant serious consideration.

A quite different method of tuning is used in the donutron, discussed in Sec. 22-12.

CHAPTER 21

OPERATING CHARACTERISTICS OF CONTINUOUS-WAVE MAGNETRONS

By GUNNAR HOK

21-1. Introduction.—For a triode or tetrode oscillator tube it is possible to obtain a set of static characteristics that can be used to calculate the performance of the oscillator under various operating conditions. With some modifications for transit-time and other effects, such an analysis is applicable also to microwave oscillators. There are two conditions that favor the use of static characteristics as a basis of oscillator theory. (1) The electronic system and the associated circuit are separate units, each of which can be studied by itself, the connection between the two being the voltage of the approximately unipotential electrodes and the currents passing through them. (2) The feedback is part of the external circuit; consequently, it can be removed and the operation of the system as an amplifier studied theoretically and experimentally.

In a magnetron oscillator these favorable conditions do not exist. A boundary surface between the electronic system and the tank circuit can be found, but the conditions at this boundary cannot in general be stated in terms of one voltage value, one current value, and the phase angle between them. Above all, the feedback cannot be removed; consequently, no information of value for analysis of the oscillating condition can be obtained by any known static test methods. The static characteristics described by Kilgore¹ and others are of no help in the study of the modern microwave magnetron oscillators. The characteristic curves and charts that are conventionally used to present the operating data of magnetron oscillators are consequently not analytical and instructive, but merely empirical and descriptive. They simply record how the performance of the oscillator in the desired mode of oscillation varies with the operating conditions. Two different charts or diagrams are obtained, depending on whether the operating condition is varied on the input side, by means of power supply and magnet, or on the output side, by means of the load.

¹ KILGORE, G. R., Magnetron Oscillators for the Generation of Frequencies between 300 and 600 Megacycles, *Proc. I.R.E.*, 24, 1140 (1936).

21-2. Variation of Performance with Input Conditions. Performance Chart.—The variables directly associated with the input are anode voltage E_b , anode current I_b , and magnetic field \mathcal{H} . Incidentally, it may seem peculiar to use the term *anode voltage*, since the anode for practical reasons is kept at ground potential and a negative voltage is applied to the cathode. These practical considerations of course are that the anode, resonant system, and transmission line are inconvenient to insulate, while the cathode and its heating transformer can without too serious disadvantages be maintained at a high voltage with respect to ground. As far as nomenclature is concerned, it is preferable to adhere to the practice adopted for multielectrode tubes of considering the cathode as the neutral reference point for the electrode voltages.

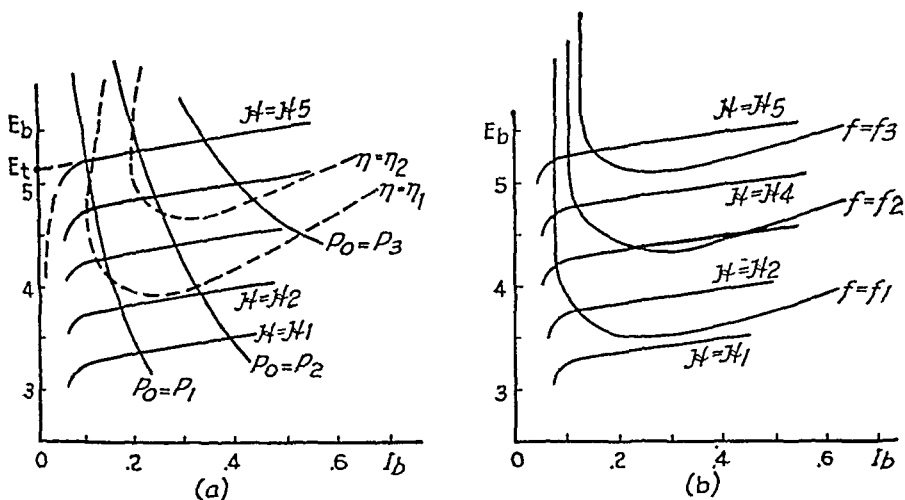


FIG. 21-1.—Magnetron performance chart.

Of the three variables mentioned, E_b , I_b , and \mathcal{H} , only two are independent. If we should select E_b and \mathcal{H} as ordinate and abscissa for a diagram of magnetron performance, we would obtain the mode-line diagram, Fig. 20-15, for the single desired mode, usually the π -mode. The line for this mode would represent the condition $I_b = 0$, and by plotting curves for other values of anode current we would get the first set of curves in this kind of performance diagram.

However, it has been found more significant to use the anode current I_b for abscissa, instead of the magnetic field \mathcal{H} (Fig. 21-1). The reason is easy to understand when we consider that one of the most important input quantities is the input power $E_b \times I_b$, which is not so simply related to the magnetic field as to the anode current.

The parametric variables that we wish to display on the chart in order to give a picture of the performance of the oscillator are the magnetic field \mathcal{H} , the output power P_o , the efficiency η , and the frequency f . These

quantities are seldom all shown on one diagram; either at least one of them is left out, or they are divided between two separate graphs, as in Fig. 21-1.

The procedure in plotting the performance chart is to mark the experimental points on the graph and write the observed values of the desired quantities close to each point. All points that represent the same value of a certain quantity are then joined by a curve, so that one set of curves is obtained for each of the dependent variables \mathcal{R} , P , η , and f .

The load impedance should remain constant throughout the measurements. For convenience a load should be used the impedance of which is to a close approximation independent of frequency within the range of the small frequency variations obtained. When the frequency readings are taken, it is advisable to eliminate frequency variations due to thermal expansion by maintaining the average power constant and sweeping along a curve of constant magnetic field by applying a low-frequency modulating voltage in series with the anode voltage. A voltage proportional to the modulating voltage is applied to the vertical deflection plates of an oscilloscope and a voltage proportional to the anode current to the horizontal deflection plates. The constant- \mathcal{R} curve is thus traced on the cathode-ray screen. A wavemeter is loosely coupled to the output, and the output of the wavemeter detector is superposed on the voltage applied to the vertical deflection plates. When the frequency of oscillation passes through the value to which the wavemeter is tuned, a small "pip" on the curve will be produced by the wavemeter output. As the tuning of the wavemeter is varied, this pip will slide along the curve, so that the frequency of oscillation can easily be determined for any point on the curve.

21-3. Curves of Constant Magnetic Field. Mode-jump Current. Figure 21-1 shows a fairly typical performance chart. The curves of constant magnetic field approach straight lines of very small positive slope. Extended to zero anode current each line intersects the voltage axis at the *threshold voltage* E_t , for that magnetic field. The first part of each curve is often impossible to observe, as the current jumps directly from zero to a finite value when the tube starts to oscillate. As mentioned in Sec. 20-8, the current reaches a maximum value as the voltage is progressively increased; if the voltage is raised beyond this point, the oscillations in the π -mode cease. In most multicavity magnetrons, oscillations will start in another mode instead. For this reason the current limit has been called the *mode-jump current*. The mode-jump current is dependent on the magnetic field and can, if low enough, be plotted on the performance chart as a demarcation line between the oscillatory and nonoscillatory regions for the π -mode. Of course the designer tries to design the tube so that this boundary line shall be far

outside the region of usual operation. Consequently, the mode-jump current is seldom indicated on the performance chart, which rarely extends beyond normal operating values of voltage and current.¹ In evaluation of magnetrons, however, particularly c-w magnetrons that are to be modulated, the mode-jump current is a very important figure which deserves special attention. The main factors that influence the mode-jump current are as follows:

1. The scaling of the electronic system, particularly the current scaling
2. The shunt impedance of the tank circuit
3. The cathode emission, but only under abnormal conditions

The first point simply states that the mode jump is an electronic phenomenon and as such subject to the scaling rules. In other words, as long as a tube is scaled from a satisfactory model magnetron it will be satisfactory also in this respect. But when a magnetron is tuned over a wide range, the scaling rules cannot possibly be followed, as pointed out in Sec. 20-8.

The shunt impedance is also affected by scaling considerations, but will be discussed separately, because it can be, and usually is, varied independently when the magnetron is tuned to a different wavelength. As a matter of fact, it is the only means available to compensate for the variations of the mode-jump current due to the change in wavelength.

Even if the electron current is never temperature-limited at any point of the cathode surface, a definite mode-jump current exists. Under normal conditions the mode-jump current is independent of the cathode emission. Whenever temperature limitation occurs, however, the mode-jump current is reduced. It is a general observation that in case of space-charge-limited current the curve for constant magnetic field approaches the drop-out point without appreciable change of slope; on the other hand, temperature limitation is revealed by a slight upward bend in the curve close to the mode-jump point.

The slope of the constant- \mathcal{H} curves is the modulation impedance. Like the mode-jump current it depends chiefly on the scaling of the tube but also on the shunt impedance. For tubes with wide tuning range the modulation impedance necessarily varies considerably with the frequency.

If for some reason a mode of lower frequency exists, or a tightly coupled spurious resonant system causes the π -mode to have more than one resonance frequency, the curves for constant magnetic field may appear as shown in Fig. 21-2. The sudden change in current and voltage is accompanied by a change in frequency, and the discontinuities in the curve occur at different points for increasing and decreasing voltage.

¹ Figure 22-3 shows a performance chart of an experimental magnetron with the curve for the mode-jump current appearing in the lower part of the diagram.

This kind of disturbance, of course, should not happen in a well-designed magnetron. The phenomenon is similar to the "long-line effect," which will be discussed in Sec. 21-8.

The general shape of the constant-output and constant-efficiency curves can be seen from the graph (Fig. 21-1a). The efficiency tends

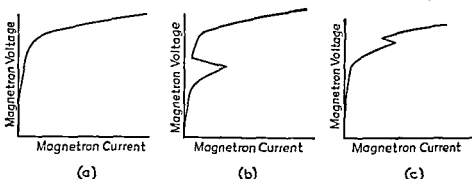


FIG. 21-2—Instability caused by spurious resonance.

to improve as the power is increased, at least if both voltage and current are increased. The efficiency usually does not vary greatly along a constant-3C curve, except for the very first part, where the current is small.

21-4. Curves of Constant Frequency. Pushing.—The constant-frequency curves are shown in Fig. 21-1b. Typical frequency variation along a line of constant magnetic field has been plotted in Fig. 21-3 in

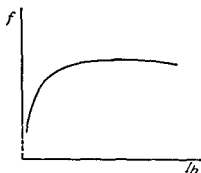


FIG. 21-3.—Pushing curve.

order to demonstrate more clearly how the frequency varies in practical operation. The fact that the frequency of oscillation changes with the anode voltage or current is usually referred to as *frequency pushing* or just *pushing*, a term coined in analogy with the previously established term *frequency pulling*, which denotes variations of frequency with load impedance. For pulsed magnetrons it is customary to use the derivative df/dE_0 at the operating point

corresponding to the relatively flat top of the pulse as a measure of the frequency pushing. For modulated c-w magnetrons the slope of the pushing curve at any one point is not very significant because of the large variations of the slope. The frequency difference between the highest and lowest points reached during modulation is of more interest.

In most applications where modulated c-w magnetrons could be used, the pushing is a serious drawback. It may therefore be of interest to

mention briefly some of the factors on which the frequency pushing in a magnetron oscillator depends.

In Sec. 20-8 the space-charge mechanism of a magnetron was described in terms of a revolving cloud of electrons, which by the action of the r-f field assumed the shape of hub and spokes of a wheel. In addition to the rotary motion of the whole system, the electrons in the spokes have a continuous drift toward the anode. For the following discussion it is convenient to assume that this wheel of space charge exists and remains unchanged as to geometry and space-charge density throughout the operation. This assumption is obviously inaccurate but good enough for a preliminary qualitative discussion. The first question is what influence the d-c plate voltage has on this revolving configuration of electrons. If the plate voltage is raised, the electric field accelerating the electrons will increase, and both the rotary motion of the cloud and the radial drift of the electrons will increase in velocity. The space-charge wheel thus behaves like the armature of a d-c motor: when the armature voltage is increased, the motor accelerates; similarly, when the magnetic field is increased, the motor slows down. The generation of r-f power in a magnetron is a process very similar to the power generation in an alternator. The rotor has electric instead of magnetic poles (the spokes of the space charge), but the frequency generated is given by the same expression

$$f = \frac{p}{2} \times \frac{s}{60} \quad (21-1)$$

where $p/2$ is half the number of poles in the alternator, or the full number of spokes in the magnetron, and s the speed in rpm.

If the plate voltage is such that the generated frequency is appreciably lower than the resonance frequency of the tank circuit, the impedance offered by the tank circuit is low and nearly a pure inductance; the power produced by the current induced by the revolving field is consequently small. The reaction on the revolving space charge from the "armature" or tank circuit is also small. This reaction is, of course, a retarding force acting on each individual electron by means of the r-f field, but it is convenient to think of this reaction as integrated over the space charges so that it forms a braking torque on the "rotor" due to the load, analogous to the torque produced by the load in an ordinary alternator.

When the speed of the space-charge wheel is increased by an increment of the plate voltage, the frequency will increase and approach resonance. The tank circuit will offer a higher impedance with smaller phase angle, so that the output power will increase rapidly with frequency. The braking torque on the space charge produced by the load increases in proportion to the output power and, consequently, tends to prevent or

minimize the changes in speed and frequency. How large the resulting frequency changes, *i.e.*, the pushing, will be depends then, as far as the tank circuit is concerned, on how steep the resonance curve is, *i.e.*, on the shunt impedance and the loaded Q of the tank circuit. As far as the space charge is concerned, the frequency change depends on the geometry and density of the space charge, *i.e.*, on the scaling of the tube and the operating conditions.

This discussion can easily be extended to explain qualitatively the mode-jump limitation. When the frequency of oscillation has reached the resonance frequency of the tank circuit, the output power and the braking torque on the space charge cannot increase any more. If the plate voltage is further increased, steady-state oscillation is not possible, because the power and torque will decrease and the space charge will rapidly gain speed until the generated frequency is so high that the imped-

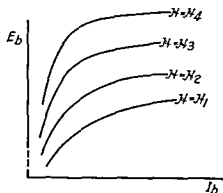


FIG. 21-4—Curves of constant magnetic field showing excessive electron leakage.

ance of the tank circuit is too small to maintain an r-f voltage sufficient for the bunching of the electrons. The oscillations will thus die out, and new oscillations can build up in another mode that happens to be favored by the operating conditions.

In this discussion it has been assumed that the geometry and density of the rotating space charge remains constant so that the device acts approximately as a constant-current generator, an assumption that within certain limits agrees with

experiments. The fact that the mode-jump limit moves toward higher power if the shunt impedance is increased is in agreement with this assumption.

Another factor that affects pushing and the performance chart in general is *electron leakage*. If electrons escape axially from the interaction space, they will take d-c power from the high-voltage supply without contributing appreciably to the r-f output. They may even absorb energy from the r-f field before hitting the wall of the end cavity and giving up their kinetic energy as heat. What happens to the performance chart can be seen from Fig. 21-4. If we follow the curves for constant magnetic field from right to left, they will start bending down toward the origin at a much larger value of anode current than the corresponding curves for a good tube. The pushing curve of Fig. 21-3 will change in the same fashion. Both efficiency and frequency stability will thus be strongly affected by electron leakage. Preventive measures are

the use of end shields of proper design and possibly a slight distortion of the magnetic field such that the electrons are "focused" toward the center of the anode.

21-5. Series Magnet.—The very small slope of the curves for constant magnetic field in the performance chart means that small voltage variations will cause very large variations in the r-f output. This disadvantage can be reduced by using a high-voltage power supply with very poor regulation or with constant-current characteristics. Another method of increasing the power stability, which has proved superior in many respects, is to derive the magnetic field from an electromagnet connected in series with the magne-

tron. The point of operation will no longer slide on a constant- \mathcal{H} line, but on a line through the origin. Figure 21-5 shows the operating characteristics of an arrangement of this kind. The oscillations start at a much lower anode voltage, and the adjustment of the anode voltage for desired output power is much less critical than with constant magnetic field. The effect of tuning on the output power is also considerably reduced. The increased power stability is bought at a price of slightly reduced frequency stability, because the line of series operation cuts the lines of constant frequency at a larger angle than the constant- \mathcal{H} lines do.

The series magnet should be provided with a variable shunt, so that an adjustment of the relation between anode current and magnetic field can be made.

It can be seen from the foregoing discussion that there is a certain analogy between a magnetron and a d-c motor. Separate, shunt, or series excitation, or any combination of them, can be used, depending on the degree of power stability and frequency stability desired for each particular application.

21-6. Variations of Magnetron Performance with Load Impedance.—For many applications of magnetron oscillators it is helpful to have a chart that shows how the performance depends on the output conditions, *i.e.*, the impedance of the load, just as the performance chart shows the dependance on the input conditions. The load impedance is complex and therefore two-dimensional. Two impedance coordinates in the

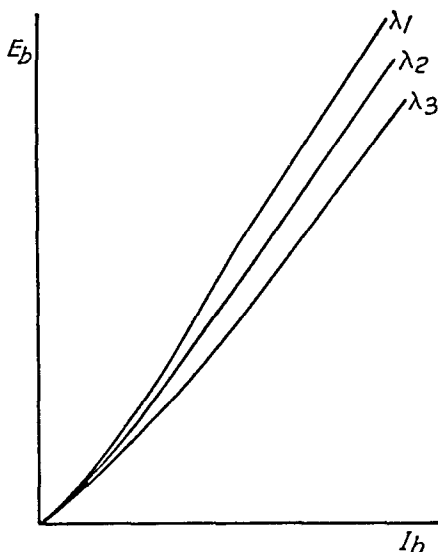


FIG. 21-5.—Volt-ampere characteristic of tunable magnetron with series magnet.

shown.¹ Here l is the distance of a minimum from the reference point, usually the output terminals, and λ_g is the wavelength (guide wavelength, if the transmission line is a waveguide). The factor 720 appears here instead of 360 because a standing-wave pattern repeats itself not every wavelength but every half wavelength. If normalized susceptance and conductance are used, instead of reactance and resistance, the coordinate system will look the same, but the order of the figures indicating position of minimum will be reversed, 0 deg appearing to the right and 180 deg to the left.

A second form of graph paper is that used in Fig. 21-7 where only concentric circles of constant standing-wave ratio and radial lines of constant position of minimum are shown. The center of the diagram represents unity standing-wave ratio, or zero reflection coefficient, while the circumference of the diagram represents infinite standing-wave ratio or unity reflection coefficient. The reflection coefficient ρ is proportional to the radius vector, and the standing-wave ratio is

$$\sigma = \frac{1 + \rho}{1 - \rho} \quad (21-2)$$

The standing-wave pattern repeats itself every half wavelength, and consequently 360 deg in the diagram is taken as $\frac{1}{2}$ wavelength. If the position of the minimum is measured from the output terminals of the magnetron, it should be plotted clockwise from the zero axis.

The normalized impedance coordinates of this diagram are shown in Fig. 2-1 usually called a *Smith chart*. A brief description of the properties, use, and interpretation of this impedance presentation² is given in Sec. 2-2, and a more detailed discussion is found in the literature. Suffice it to say here that the reader should be familiar with the identity of the two representations of Figs. 2-1 and 21-7. The latter has the advantage that standing-wave data can be entered directly without computation or graphical construction; the former facilitates the interpretation of the data in terms of reactance and resistance, or susceptance and conductance, at the reference point.

21-7. The Rieke Diagram.—The standard chart adopted to record the variation in performance of a magnetron with load impedance is generally called the *Rieke diagram*. It may use the coordinate system of

¹ The fact that the standing-wave ratio is indicated as the square root of a number is a matter of convenience explained by the circumstance that most standing-wave indicators use square-law detectors such as crystals, thermocouples, bolometers, etc. The number under the square-root sign is thus the ratio of the maximum to the minimum microammeter deflection (unfortunately often referred to as the *power-standing-wave ratio*).

² SMITH, PHILLIP H., Transmission Line Calculator, *Electronics*, 12, 29 (1939); 17, 130 (1944).

either Fig. 2-1 or 21-7, for practical reasons usually the latter. One of the great advantages of this representation is that the shape of any curve, *e.g.*, the curve of constant output power vs. load impedance, is practically independent of the position of the reference point, as long as the frequency is approximately constant. A change of reference point only rotates the whole diagram through an angle of 720 deg times the

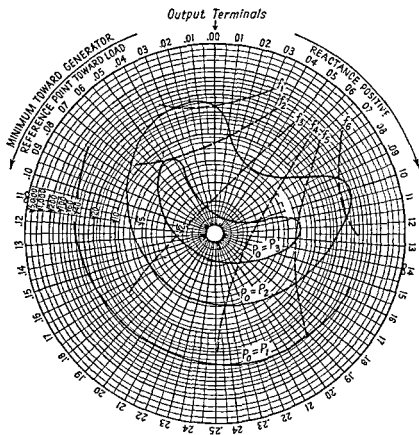


FIG. 21-7.—Rieke diagram. Radial coordinate, $\frac{\sigma - 1}{\sigma + 1}$ where σ is the standing-wave ratio; azimuthal coordinate, l/λ_g ; l , distance between output terminals and standing-wave minimum (clockwise direction positive).

ratio of the distance between the old and the new reference point to the wavelength. If, for example, the reference point is moved from the output terminals of the magnetron $\frac{1}{8}$ wavelength along the line toward the load, the diagram will be rotated counterclockwise through an angle of 90 deg.

Figure 21-7 shows a typical Rieke diagram for a magnetron. In order to show more clearly the fundamental shape of the constant-power and constant-frequency curves, a typical plot in Cartesian coordinates, of

normalized load susceptance vs. normalized load conductance, is shown in Fig. 21-8. In this diagram the reference point has also been changed; the abscissas of the curves are not the values of the load admittance at the output terminals but the load admittance at the point *BB* in the equivalent circuit of Fig. 20-13. The load thus appears directly across the normalized tank circuit, as shown in Fig. 21-9. If the normalized space-charge admittance were a pure negative conductance and dependent only on the amplitude of oscillation, the curves of constant power in Fig. 21-8 would be expected to be straight lines of constant load conductance, because in any steady state the space-charge conductance must necessarily be equal to the negative of the sum of the conductance of the unloaded tank circuit and the conductance of the load. The frequency of oscillation would be that

for which the tank-circuit susceptance is equal to the negative of the load susceptance; consequently, the constant-frequency curves should

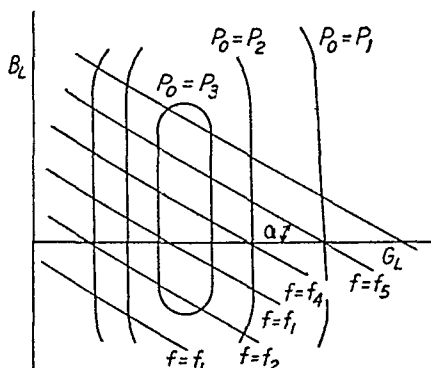


FIG. 21-8.—Curves of constant output and frequency in Cartesian admittance plane.

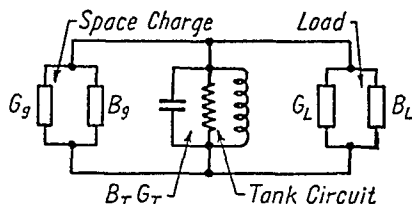


FIG. 21-9.—Simplified equivalent circuit of magnetron oscillator.

be straight lines of constant load susceptance. When the load susceptance is not too great, the constant-output curves in Fig. 21-8 have the expected form, but the power falls off at large susceptance values, probably because of the fact that the field pattern loses its symmetry and regularity. The presence of other modes than the π -mode in the r-f field reduces the efficiency of conversion in the space-charge cloud.

The constant-frequency curves are straight lines, not parallel to the conductance axis but intersecting this axis at a constant angle α . In other words, the equivalent susceptance representing the tuning effect of the rotating space charge at constant anode current is not constant but varies linearly with the conductance of the load.¹

¹ The significance of the angle α has not been quantitatively explained by magnetron theory, but a certain similarity between klystron oscillators and magnetron oscillators gives an idea of what factors determine α . For a klystron oscillator the constant-frequency curves have the same character as those for the magnetron, and the theory gives a simple relation between the angle α and the transit angle between buncher and catcher. In a magnetron each gap between two anode segments acts as a buncher for some electrons and as a catcher for other electrons. The transit

In the Rieke diagram of Fig. 21-7 the constant-output curves coincide with constant-conductance circles in the central parts, and as the coordinates of this chart are also orthogonal, the constant-frequency lines will intersect the constant-susceptance circles at the same angle α . The point of maximum output power is called the *power center* and the region where the constant-frequency curves converge, the *frequency sink*.

The anode voltage and, consequently, the input power do not vary greatly over the central parts of the diagram; therefore the constant-efficiency curves follow the constant-output curves fairly closely and are usually left out of the diagram altogether.

The importance of the Rieke diagram is that it gives the equipment designer information about the optimum load conditions. For a well-designed tube the power center should be fairly close to the center of the diagram, i.e., unity standing-wave ratio. In many applications, however, the standing-wave ratio will vary within certain limits, and the equipment designer has to know how this will affect the operation of the oscillator. In the high-conductance part of the diagram, i.e., the region around the frequency sink, there may be a nonoscillatory region due to the mode-jump condition described in connection with the performance chart, and the designer has to make sure that this region will never be reached in actual operation.

Frequency pulling, i.e., the variation of frequency with load impedance, is also an important factor in many applications. The constant-frequency curves give information about this phenomenon. Close to the frequency sink, where these lines crowd together, the frequency stability is poor. The conventional measure of the pulling of a magnetron—or other self-excited oscillator—is the *pulling figure*, which is defined as the difference in megacycles between the maximum and minimum frequencies obtained when the load impedance is varied in such a way that the standing-wave ratio remains constant and equal to 1.5 and a voltage minimum moves continuously at least $\frac{1}{2}$ wavelength, i.e., that the point representing the load impedance in the Rieke diagram describes a complete circle with the radius equal to $(1.5 - 1)/(1.5 + 1)$, or 0.2.

The pulling figure is related to the external Q , Q_s , of the tank circuit and to the angle α between the constant-susceptance and constant-frequency lines in the Rieke diagram. This relation is easily derived from the diagram in Fig. 21-8. The normalized tank-circuit susceptance seen from the output terminals is $2\delta Q_s$, where δ is $(\omega - \omega_0)/\omega_0$, ω_0 is 2π times the resonance frequency f_0 of the tank circuit, and ω is 2π times

angle between gaps is determined by the geometry and the angular velocity of the electrons, which depends on anode voltage and magnetic field. If all other conditions remain constant, the transit angle and α would thus be expected to decrease with increased anode voltage.

the frequency of oscillation f . The locus of admittance with the standing-wave ratio σ equal to 1.5 is a circle, the diameter of which is equal to the distance between its intersections with the horizontal axis which is $\sigma - 1/\sigma$. Thus we have

$$\frac{\sigma - \frac{1}{\sigma}}{\cos \alpha} = 2(\delta_1 - \delta_2)Q_E \quad (21-3)$$

where δ_1 and δ_2 are the maximum relative frequency deviations, positive and negative, from the tank-circuit resonance. The pulling figure

$$\begin{aligned} (\delta_1 - \delta_2)f_0 &= \frac{f_0}{2Q_E \cos \alpha} \left(1.5 - \frac{1}{1.5}\right) \\ &= \frac{5}{12} \frac{f_0}{Q_E \cos \alpha} \end{aligned} \quad (21-4)$$

It is assumed that the changes of f_0 and Q_E caused by the presence of the space charge are small enough to be neglected in this expression.

The fact that the constant-frequency curves in the Rieke diagram depend on the external Q can be used to determine this quantity in the following manner. The frequency difference $f_1 - f_2$ between two points on the same constant-output curve close to the center of the diagram is measured. Since the curve is one of constant power output, the normalized load conductance will be constant. Values of the normalized load susceptance at these points, B_{L1} and B_{L2} , are read from the graph if these coordinates are shown; otherwise they are calculated. Then

$$B_{L1} - B_{L2} = 2(\delta_1 - \delta_2)Q_E$$

and consequently

$$Q_E = \frac{B_{L1} - B_{L2}}{f_1 - f_2} \frac{f_0}{2} \quad (12-5)$$

where

$$\delta_1 = \frac{f_1 - f_0}{f_0} \quad \text{and} \quad \delta_2 = \frac{f_2 - f_0}{f_0}$$

By selecting points on the same constant-power curve we avoid the necessity of correcting for the angle α .

12-8. Long-line Effect.—Some interesting complications arise from the fact that a transmission line is a resonant system when the standing-wave ratio is greater than unity. It has long been observed that self-excited oscillators are unstable when loaded by a resonant load. If the coupling between the tank circuit and the load exceeds the critical coupling, and the load is tuned to resonance with the tank circuit, more than one possible frequency of oscillation exists, and the oscillator may shift from one to the other at the slightest provocation.

In order to analyze this phenomenon for a magnetron that is connected to its load by a transmission line, we plot the susceptances of the load and the tank circuit, respectively, vs. frequency, as shown in Fig. 21-10. For the time being it will be assumed that the susceptance of the space charge is negligible ($\alpha \approx 0$); it will be seen later that the diagram can be revised slightly so as to take the space-charge susceptance into account. Note that the load susceptance has been plotted with reversed sign, so that any intersection between the two curves means that $B_{\text{load}} + B_{\text{tank}} = 0$, and the corresponding frequency may be a possible frequency of oscillation. For the small frequency band close to resonance considered here, the tank-circuit susceptance curve is a straight line with a slope proportional to the external Q . The load susceptance in the diagram has been obtained under the assumption that the line is terminated in a resistance

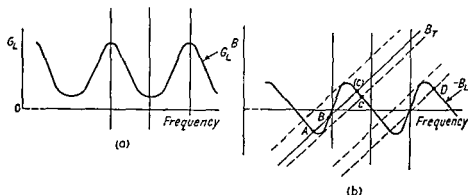


FIG. 21-10.—Long-line effect in tunable magnetron G_L , load conductance; B_T , tank susceptance; B_L , load susceptance

or an impedance that can be represented by a certain length of line and a resistance. The standing-wave ratio is thus constant within the small frequency range considered. The load susceptance varies periodically between two limits $\pm B_{\text{max}}$, faster when the point representing the admittance passes through the high-conductance region of the Rieke diagram and somewhat more slowly on the low-conductance side of the diagram. The load conductance, as shown in Fig. 21-10, varies periodically between a maximum and a minimum. It is included to show which of the intersections between the susceptance curves represent admittance minimums or parallel resonance. Only these give possible frequencies of oscillation for an oscillator of the negative-conductance or shunt type considered here. Because the negative conductance of the space charge can be varied from zero to a certain maximum, the frequencies of parallel resonance or admittance minimums are favored. For a series-type oscillator such as an arc discharge, having a negative resistance that could be varied between zero and a certain limit, the impedance minimums (or

admittance maximums) of the combined tank and load circuits would give the preferred frequencies of oscillation. In Fig. 21-10 then the points *A* and *C* represent possible frequencies of oscillation for a magnetron, while the point *B* does not.

Suppose that the frequency of oscillation corresponds to point *A* and that the tank circuit is tuned gradually to higher frequencies. In Fig. 21-10 the line B_T representing the tank-circuit susceptance will move toward the right, parallel to itself. The intersections *A* and *B* will move toward each other and finally coincide when B_T becomes tangent to the B_L curve. If the frequency is increased further, the tank susceptance line will no longer touch the *AB* lobe of the load-susceptance curve. The oscillations cease, and new oscillations build up with a frequency corresponding to the point *C*. The frequency will jump from one lobe of the curve to the next one to the right through the full tuning range

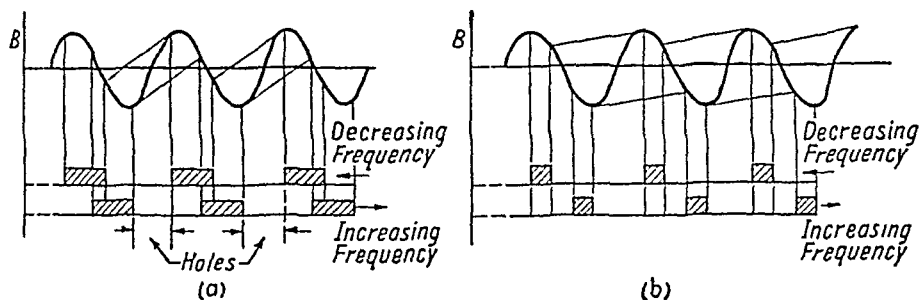


FIG. 21-11.—Reduction of frequency coverage by long-line effect.

of the tube. If the direction of tuning is reversed, the same phenomenon will be repeated, but the oscillation will now jump from a point near the upper crest of the lobes to a point on the sloping side of the lobe next to the left.

One of the important consequences of this so-called *long-line effect* is that continuous coverage is not obtained over the tuning range of the oscillator. Figure 21-11 shows the frequency bands covered for both directions of tuning. In order to fill in the holes in the tuning range, it is necessary to eliminate the long-line effect or to provide means for an adjustment of the length of the transmission line. The latter process is usually carried out by means of a so-called *line stretcher* of some sort, e.g., a telescoping "trombone" section of line. In a waveguide the variation of the electrical length of the line can be conveniently accomplished by moving a long tapered dielectric slug from one side of the guide in the direction of the \mathcal{H} vector toward the center of the guide and thereby changing the phase velocity in that part of the guide. The result of varying the length of the line is very much the same as the effect of varying the tank-circuit tuning, but the $-B_L$ curve will slide toward the right

or the left, instead of the B_T curve. The long-line effect may, therefore, interfere with the procedure of obtaining data for a Rieke diagram. If readings are being taken at constant standing-wave ratio, *i.e.*, following a circle around the center of the diagram (Fig. 21-12) a point F may be reached where the frequency suddenly shifts and the new admittance is found to correspond to the point G . If the line stretcher is moved in the opposite direction, the corresponding points F' and G' are found. It is thus found that there is a region in the Rieke diagram where measurements cannot be made. This is referred to under various names, such as *instability region* or *forbidden region*. This region should not be confused with the region defined by the mode-jump current, which has been discussed earlier.

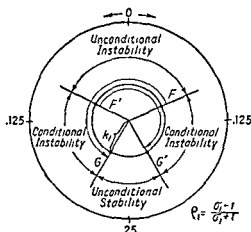


FIG. 21-12.—Example of the long-line effect plotted on the Rieke diagram for given standing-wave ratio σ_1 .

In the case of the mode-jump phenomenon the oscillations cease or switch to a different mode. The boundaries of this region are characteristic of the tube and rightly belong in the Rieke diagram. The instability region, on the other hand, is in addition dependent on the resonant properties of the load and should therefore not be shown in the Rieke diagram, which should be quite general and independent of the test equipment used.

Before various methods to minimize or eliminate the long-line effect are discussed, it may be worth while to investigate to what extent the presentation of the long-line effect will have to be modified to take into account the influence of the space charge on the frequency of oscillation. In the equivalent circuit of Fig. 21-9 the normalized space-charge susceptance can be thought of as made up of two components: B_{g0} , which is independent of the load conductance, and B_{g2} , which is proportional to the load conductance, actually equal to $\tan \alpha$ times the load conduct-

ance. The angle α is as before the angle between the constant-frequency and constant-susceptance curves in the Rieke diagram. It is apparent from Fig. 21-8 that this relation is true. The susceptance condition for oscillation can now be written

$$B_T + B_{g0} = -(B_L + G_L \tan \alpha) \quad (21-6)$$

instead of

$$B_T = -B_L \quad (21-7)$$

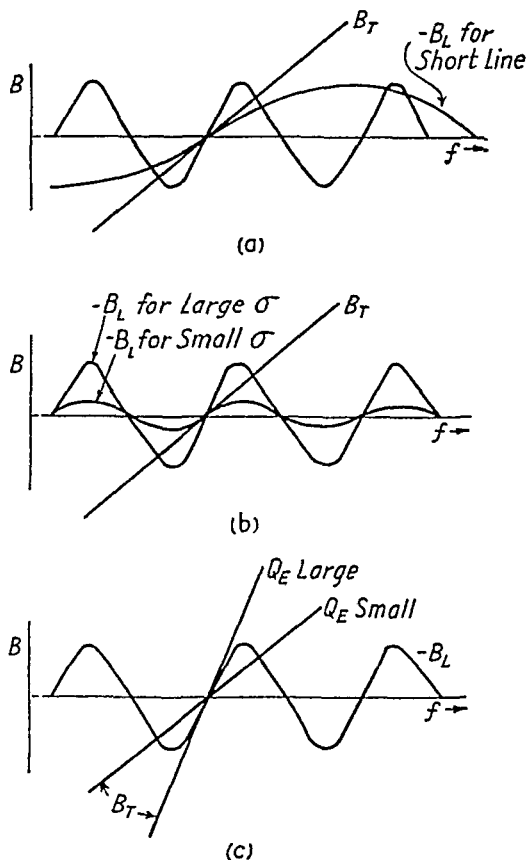


FIG. 21-13.—Elimination of the long-line effect by (a) reduced line length; (b) reduced standing-wave ratio; (c) increased external Q .

For a small frequency range the left-hand side can still be represented by a straight line, and the right-hand side by a wavy curve the phase position of which has been changed in relation to the simple plot in Fig. 21-10. Actually this diagram could just as well represent the graphical solution of Eq. (21-6) as of Eq. (21-7). The qualitative discussion of the long-line effect is thus entirely unaffected by the simplifying assumption.

If the transmission line is a waveguide, there will be another slight change in the $-B_L$ and G_L curves of Fig. 21-10. Because of the fact that the phase velocity in the guide decreases with increasing frequency, the "wavelength" of these curves along the frequency axis will increase with increasing frequency. The characteristic impedance of the guide also varies with the frequency; if the line is terminated by a constant resistance, the standing-wave ratio and the amplitude of the $-B_L$ and G_L curves will therefore change somewhat with frequency. The qualitative presentation of the long-line effect, however, is perfectly valid also for a magnetron coupled to a waveguide.

Quantitatively, the important parameters for the long-line effect are the "wavelength" of the load-susceptance curve along the frequency axis, which is equal to the fundamental frequency of the transmission line, and therefore depends upon the length of the line; the standing-wave ratio σ , which determines the amplitude of this curve, $(\sigma^2 - 1)/2\sigma$ and the external Q of the tank circuit, which determines the slope of the straight line representing the tank-circuit susceptance, this slope being $2Q_E/f_0$ where Q_E is the external Q and f_0 the resonance frequency of the tank circuit. As indicated in Fig. 21-13 any one of these three parameters can be used to eliminate the long-line effect. If the line is shortened, the frequency separation between the line resonances is increased, i.e., the "wavelength" along the frequency axis in the diagram becomes greater. The long-line effect disappears entirely when the slope of the wavy $-B_L$ curve at no point exceeds the slope of the B_T curve. The same result is obtained if the standing-wave ratio and therefore the $-B_L$ amplitude is reduced. Obviously a sufficiently increased slope of the B_T line will also remove the multiple intersections between the curves. If the length of the line and the maximum standing-wave ratio are determined by the specific application for which the tube is to be used, the last is the only remaining possible method for elimination of the long-line effect. To increase the slope $2Q_E/f_0$ is the same as to reduce the coupling between the tank circuit and the load. In terms of the lumped-constant-circuit theory, the coupling is reduced to or below the critical coupling. The coefficient of coupling at critical coupling is $1/\sqrt{Q_1 Q_2}$, where Q_1 of the primary circuit is the unloaded Q of the tank circuit and Q_2 of the secondary is the effective Q of the load. The reduced coupling will in most cases result in lower efficiency, electronic efficiency as well as circuit efficiency, and the power center in the Rieke diagram will move toward the high-conductance part of the diagram.

Figure 21-14 shows qualitatively how the instability region changes with the length of line or the external Q . It should be noted that these two ways of varying the size of the instability region are by no means equivalent. When the line length alone is changed, the curves of con-

stant power and of constant frequency in the Rieke diagram remain unaffected; when the external Q is changed, they do not.

The lines in Fig. 21-14 mark the boundaries of the *unconditional instability region* i.e., the region the oscillations avoid unconditionally. On both sides of this area there are zones of *conditional instability*, where the point of operation will remain only if brought there from a certain direction. Only if there is an interval where the curves B_T and $-B_L$ in

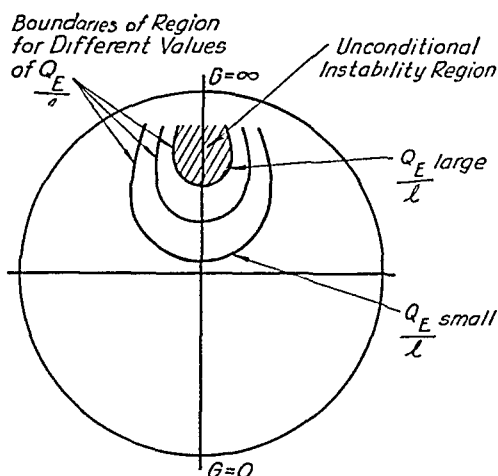


FIG. 21-14.—Region of unconditional instability for different loads.

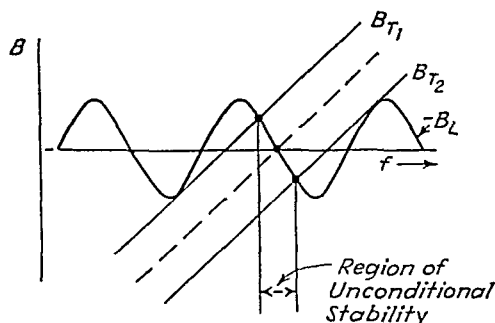


FIG. 21-15.—Region of unconditional stability, Δf .

Fig. 21-10 have just one point of intersection (Fig. 21-15) does a zone of *unconditional stability* exist. It is, therefore, often necessary to operate in a region of conditional instability. Once oscillation has been established and as long as it is maintained, this operation does not present any difficulties. If oscillation is interrupted and restarted, however, it may very well come back on a frequency corresponding to another intersection of the B_T and $-B_L$ curves. The space charge has approximately the same ability to support any one of the possible oscillations, because the field distribution in the interaction space is the same for all of them and the difference in frequency between them is very small. The load con-

ductance, on the other hand, may differ substantially, and the oscillation will occur at the frequency for which the load is lightest. Including the effect of both the space charge and the network we can say that those oscillations will gain the upper hand which have the largest initial "negative decrement" or "logarithmic increment," if the building up of oscillation in a negative-resistance system can be considered analogous to the decay of oscillation in a positive-resistance system.

21-9. Modulation.—From the magnetron operating characteristics discussed in the previous sections of this chapter, some conclusions can be drawn about the properties of the magnetron as a modulated oscillator. In the present section, these results will be compared with experimental evidence.

The performance chart gives us information of the modulation produced by changing the anode voltage at a slow rate while the load impedance is held constant. It is evident that the modulator power required is only a small fraction of the input power to the magnetron, because the slope of the curves for constant magnetic field is relatively small. The modulation is subject to a certain amount of nonlinear distortion, because the anode current varies linearly with the anode voltage and the r-f output voltage is approximately proportional to the square root of the anode current. Various modulation tests have been reported to give 20 to 30 per cent distortion with a split-anode magnetron at 80 per cent modulation and less than 10 per cent distortion with a multicavity magnetron at 100 per cent modulation.

There may be other causes of serious distortion in a modulated magnetron than this nonlinear but continuous relation between modulating voltage and r-f output voltage. If the current at the point of operation exceeds the mode-jump limit, the oscillations will be suddenly interrupted or shift to another frequency and may restart or shift back to the original frequency just as suddenly when the operating point has been brought back into the oscillatory region of the π -mode. The oscillations may also stop and start abruptly at some point at the low-current end of the modulation cycle. The transients occurring when the oscillations decay and build up at these discontinuities give high-order side bands and may cause interference in adjacent communication channels. For these reasons it is important that the mode-jump current of the tube be high enough to accommodate the largest modulation peaks without mode jump and that the tube oscillate smoothly as far down on the curve as the inverse modulation peaks will ever reach.

So far we have dealt only with slow modulation, for which the steady-state conditions recorded in a performance chart apply without modification. If, on the contrary, the ratio of the modulation frequency to the frequency of oscillation is not negligible in comparison with the reciprocal

of the loaded Q of the tank circuit, the modulating process will be somewhat different. Let the oscillator be represented by a negative conductance $-G$, a tank circuit LC , and a load G_L , as shown in Fig. 21-16a, and the susceptance of load and space charge be neglected for the time being. The oscillator produces an r-f voltage E_L across the load. Experimental data (performance charts, Rieke diagrams, etc.) indicate that the negative conductance at constant magnetic field varies with the plate voltage E_b and the r-f voltage E_L approximately as shown in Fig. 21-16b. At a certain plate voltage E_{b2} the steady state is indicated by the point P_1 where $G = G_L$. Now suppose that the plate voltage is abruptly changed to E_{b1} . The energy storage in the tank circuit cannot change discontinuously, so that in the first instant the point of operation follows a line of

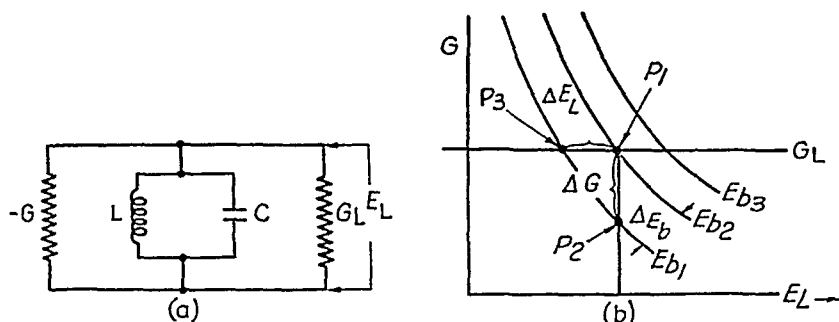


FIG. 21-16.—Analysis of magnetron modulation: (a) equivalent circuit; (b) result of sudden change of anode voltage.

constant E_L to the point P_2 . Consequently, there will be initially a net conductance ΔG in the system. The oscillations will start to decay as free oscillations in an oscillatory circuit with an effective Q of

$$Q = \frac{1}{\Delta G \omega L} \quad (21-8)$$

The unusual feature about this effective Q , compared with the Q of linear passive networks, is that it depends on how large the change in anode voltage ΔE_b is. Furthermore, it increases gradually as the oscillator approaches the new steady state, which is represented by the point P_3 in the diagram. This effective Q is always larger than the loaded Q of the tank circuit. We should expect this variable Q to have a discriminating effect against high modulating frequencies and small-amplitude modulation components. Thus while the modulation band width of a modulated amplifier is unambiguously given by f_0/Q_L , the band width of a modulated magnetron, or other self-excited oscillator, is smaller than f_0/Q_L and varies with the amplitude of the modulation.

The appearance of the r-f spectrum of the output of a modulated magnetron is largely determined by the frequency modulation caused by

pushing. Therefore sinusoidal modulation with a relatively low frequency will give, not the discrete carrier and the pair of first-order side bands, plus some small higher order side bands due to nonlinear distortion that we expect for an r-f amplitude-modulated wave, but rather a broad frequency-modulation spectrum with a very large number of side bands, as shown in Fig. 21-17. The only indication of the presence of amplitude modulation is the lack of symmetry of the spectrum.

The question then arises whether this frequency modulation by pushing can be utilized to advantage for transmission of intelligence. The

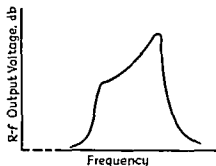


FIG. 21-17.—R-f spectrum of modulated magnetron.

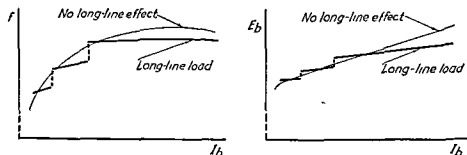


FIG. 21-18.—Effect of a long line on curves of anode voltage and frequency vs. anode current.

poor frequency stability of the carrier wave can of course be improved to any desired degree by means of automatic frequency control. The prospects of obtaining the linearity of modulation required for high-quality transmission are not so good as could be desired, however. The pushing curve in Fig. 21-3 shows how the frequency of oscillation varies with the anode current in a typical magnetron. It is far from linear; its slope in many magnetrons even changes sign not far from the high-current end of the curve.

Other serious difficulties in the use of a modulated magnetron for transmission of voice or video intelligence arise from the long-line effect. The preceding discussion of magnetron modulation was based on the performance chart, which is determined for a constant load impedance,

independent of the frequency. If, instead, the magnetron is working into a frequency-sensitive load, particularly a long, unmatched line, the operation will be radically different. Figure 21-18 shows qualitatively how the anode voltage and the frequency will vary with the anode current under these conditions. The modulating voltage changes the space-charge susceptance and thus produces approximately the same effect as a variation in the tuning of the tank circuit, *i.e.*, a movement of the B_T curve in Fig. 21-10, parallel to itself along the frequency axis. The point of operation will jump from lobe to lobe of the $-B_L$ curve in the manner described in the preceding section. Each jump corresponds to a discontinuity in the curves of voltage and frequency in Fig. 21-18. These discontinuities will cause serious distortion of the amplitude and frequency modulation and produce undesirable transient effects. The remedies are those already mentioned: the use of a short line, high external Q , or perfectly matched load, all rather difficult to obtain without sacrifice of other desirable characteristics, such as efficiency and simplicity of operation.

CHAPTER 22

DESCRIPTION, CHARACTERISTICS, AND OPERATION OF TYPICAL CAVITY MAGNETRONS

By G. Hox, H. W. Welch, and W. G. Dow

22-1. Introduction.—In this chapter some representative magnetrons with internal resonators will be described and their operating characteristics briefly reviewed. In order to give a somewhat complete picture of the state of the art as far as c-w magnetrons are concerned, it has not been possible to limit this discussion to tubes already in production or ready to go into production, although most emphasis will be placed on such tubes.

In the introduction to Chap. 20, the c-w magnetrons were classified according to power output; of the classes mentioned we find magnetrons with internal resonators in the 10-kw class, the 1-kw class, and the 50-to-100-watt class.

The 10-kw class as a whole is still in the experimental stage, but the model 34 (Fig. 22-1) magnetron has a number of interesting features and has proved beyond any doubt that magnetrons capable of generating c-w power of the order of 20 kw at 500 Mc or higher can be built. A brief description of this tube is therefore well justified.

In the 1-kw class about two-thirds of the frequency range 500 to 10,000 Mc has been covered by a sequence of experimental tubes. Only one of them, the 6J21, 2400 to 3600 Mc, is in small-scale production. All the tubes were multivane magnetrons scaled from the same model. A standard anode voltage of 2000 volts was adopted for the lower frequency tubes, while 4000 to 5000 volts was considered more suitable for the higher frequency tubes such as the 6J21. After the oxide-coated cathodes had been abandoned in favor of pure-tungsten and thoriated-tungsten filaments, it was found that all tubes, without rescaling, could be operated to advantage at an anode voltage of 4000 volts.

Unlike the 1-kw multicavity tubes and the 150-watt split-anode tubes, the low-power tubes were initiated to meet an immediate need that could not wait for the time when the higher power tubes would be available, or for special applications for which the latter would be too heavy or in other respects unsuitable.

The 4J60 to 4J65 series, which covers a total range of 2230 to 4030 Mc, falls in this category and will be discussed below.

The drawback of this series is the small tuning range of the individual

tubes, about 10 to 12 per cent. A nearly prohibitive number of tubes is therefore required to cover any considerable frequency range.

The "donutron" stands in a class by itself, although its power output is of the same order of magnitude as the 4J60. It promised wide tuning range and structural simplicity but required more time to develop, because it is definitely off the beaten path in magnetron design. Even though it is not ready for production, it holds so much promise for the future that it cannot be omitted here.

22-2. Model 34 Magnetron.—The model 34 magnetron is designed to operate at an anode voltage of 10 to 14 kv with an average modulated output power of 10 kw over the frequency range of 350 to 500 Mc.

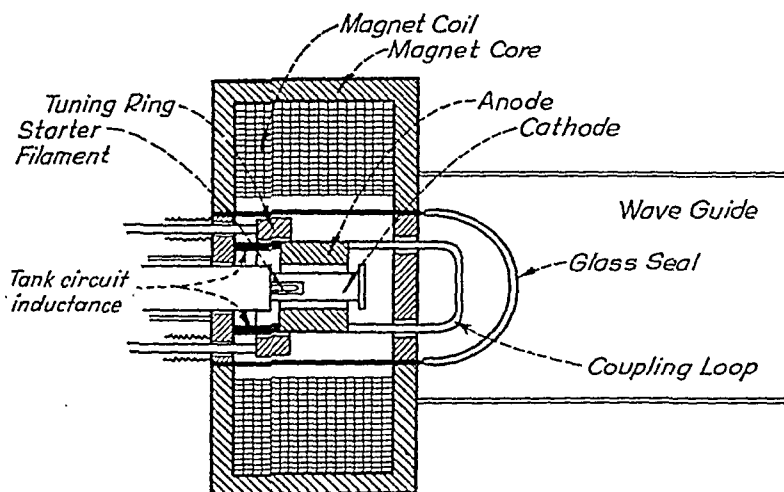


FIG. 22-1.—Schematic representation of the model 34 10-kw magnetron.

Figure 22-1 shows schematically the arrangement of the vital parts of the 10-kw magnetron. The cylindrical cathode is coated with a secondary emitter (Dowmetal in the final tube). A hairpin tungsten filament supplies the electrons required to start the oscillations and the secondary emission from the main cathode. The filament is placed in a narrow slot in the cathode so as to be protected as well as possible from electron bombardment. The tubular anode, coaxial with the cathode, is split into two equal segments. The electrodes are placed along the axis of a solenoid magnet. The iron case enclosing the coil is continued by two parallel steel plates in the magnetron itself, which form the pole pieces for the uniform axial magnetic field in which the electrodes are immersed. The cylindrical steel shell of the tube is so thin that it does not form a substantial magnetic shunt across the gap. The tank-circuit inductance consists of two copper pipes that join the anode segments to one of the copper-plated steel plates (the left one in Fig. 22-1).

The frequency is varied by means of a movable tuning ring, which increases the capacitance between the anode segments when moved to the right and reduces the tank-circuit inductance when moved to the left. We thus have in effect a combination of L -ring and C -ring tuning. A high-impedance coupling loop is connected between the anode segments in parallel with the tank-circuit inductance. It passes through the steel plate to the right and reaches into the otherwise short-circuited end of a waveguide and transfers the generated r-f power to the waveguide. In the waveguide the coupling loop is widened to a broad strip in order to keep as small as possible the self-inductance of the loop, *i.e.*, the inductance not linked with the magnetic field of the dominant mode in the waveguide.

Cooling systems and structural details have been left out or simplified in the figure. Actually the anode, the cathode, the tank-circuit inductance, the coupling loop, and the tuning ring are water-cooled, while the magnet and the output seal are cooled by compressed air.

One of the difficulties encountered in the development of this tube was overheating of the output seal, presumably by bombardment of electrons. This condition was accompanied by a glow or fluorescence at the seal. As the most elaborate traps to prevent electrons from leaking through from the interaction space did not eliminate the trouble, it was concluded that at least the greater part of the participating electrons were generated by secondary emission at the glass itself. A very thin high-resistance layer of gold was finally applied on the inside of the glass to reduce the secondary emission and to accumulate a negative charge that would repel electrons. This proved to be effective in protecting the seal.

The water-cooled secondary-emission cathode is one of the outstanding features of this tube. The starter filament can build up the revolving space charge under nonoscillatory conditions, and, when oscillation starts, the bombardment of the cathode by electrons and the resulting secondary emission set in simultaneously. A number of secondary emitters have been tried, such as beryllium-copper, aluminum, and Dowmetal. The last one is the most satisfactory emitter, but some difficulties have been experienced in establishing a coat of Dowmetal that will remain permanently on the copper cylinder. If the thermal contact between the coating and the copper is not satisfactory at all points, the Dowmetal may overheat and evaporate, leaving bare spots with negligible emission. When aluminum was used in the early models, satisfactory emission was obtained only for a short time—a matter of a few hours—after exhausting the tube. The cathode could be reactivated by letting in air and exhausting the tube again. Evidently the secondary emission was connected with the presence of gas on the surface of the cathode. The cleaning up incident to operation left the cathode stripped of its emitting properties.

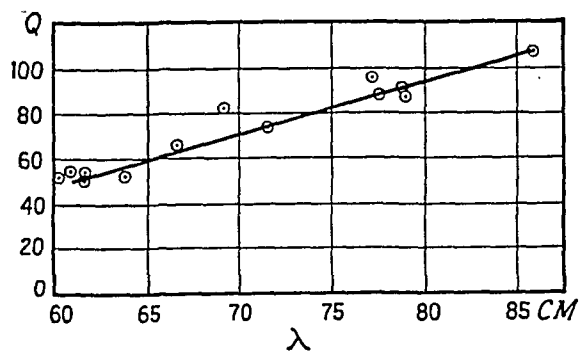


FIG. 22-2.—External Q of model 34L.

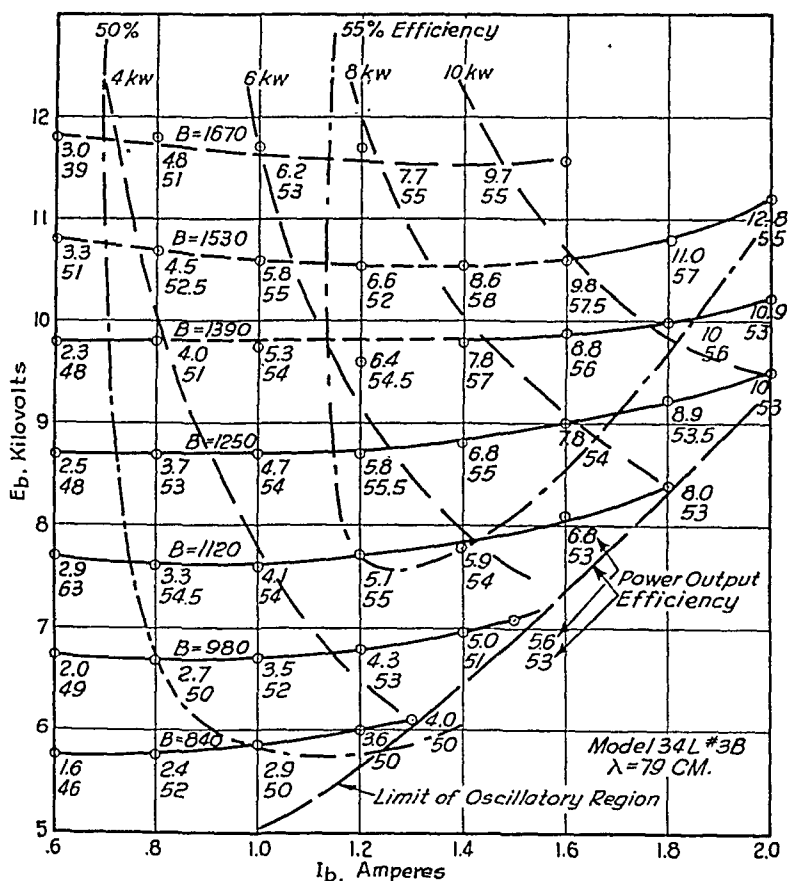


FIG. 22-3.—Performance chart for model 34L.

Figure 22-2 shows the variation of the external Q over the tuning range, obtained from data for a number of model 34 tubes. Increase of external Q with the wavelength is desirable from several viewpoints. The cavity losses are larger at the short-wave end of the range, and a low Q is required there to maintain a reasonable circuit efficiency. Furthermore, to obtain a sufficient mode-jump current requires a higher shunt impedance at the long-wave end. As $\sqrt{L/C}$ does not vary appreciably when combined L - and C -ring tuning is used, the Q variation shown answers the demand for a shunt impedance increasing with wavelength. The variation obtained is chiefly due to the fact that the characteristic impedance of the waveguide increases as the cutoff wavelength is approached. For the 22- by 6-in. waveguide used, the cutoff wavelength is 112 cm. The undesirable effect of the increased Q is that the maximum usable modulation band width is reduced at the long-wave end of the tuning range.

Figure 22-3 shows a performance chart for a model 34 magnetron at 79-cm wavelength. Even higher output power and better efficiency have been obtained. When a power of 19.5 kw was measured, there was no indication that the tube characteristics prevented further increase. The available power supply, however, had reached the limit of its capacity. The efficiency in the high-power region of the chart varies from 55 per cent to slightly over 70 per cent, according to wavelength and operating conditions. There were often indications of a negative slope at the low-current end of the curves for constant magnetic field. This tends to cause relaxation oscillations, which break the r-f output up into pulses. In the performance chart of Fig. 22-3 dashes in the curves that pulsing has interfered with the taking of data. Typical operating data are given in Table 22-1.

TABLE 22-1—OPERATING DATA OF MODEL 34P MAGNETRON

Starter filament voltage	2 4 volts
Starter filament current	65-70 amp
Frequency range.	325-495 Mc
Typical Operation	
Frequency	490 Mc
Filament current	68 amp
Magnetic field	1670 gauss
Anode voltage	12.4 kv
Anode current.	1.4 amp
Power output.	11.5 kw
Efficiency . . .	66%

22-3. The 6J21 1-kw Tunable Magnetron.—The 6J21 is a vane-type magnetron that is tunable over the range from 2460 to 3610 Mc. The output power is rated as greater than 1000 watts between the limits 2720

to 3150 Mc. Maximum plate dissipation is 1000 watts. Both air and water cooling are required.

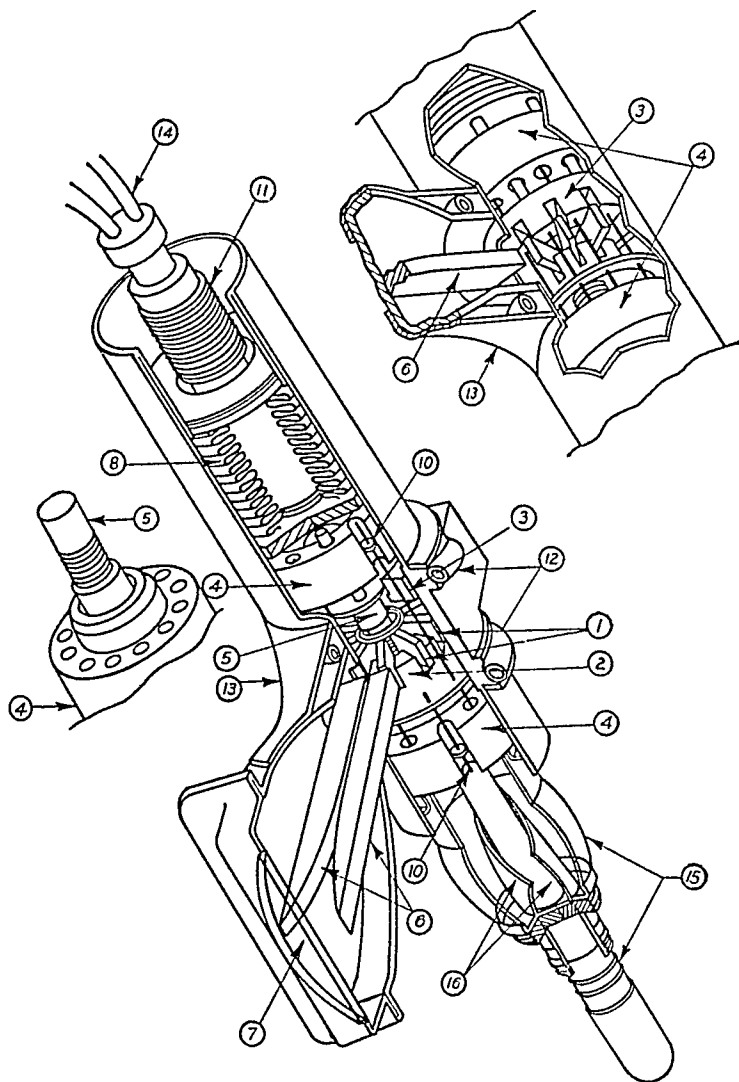


FIG. 22-4.—Cutaway view of 6J21 magnetron: (1) vanes; (2) C-ring plunger; (3) L-ring plunger; (4) magnet poles; (5) filament end hat; (6) output transformer ramps; (7) output aperture; (8) tuning bellows; (9) tungsten guiding rods for tuning assembly; (10) copper bearing supports for guide rods; (11) threaded post for tuning nut; (12) water cooling pipes for anode; (13) magnet "shoe"; (14) inlet for tuner cooling water; (15) cathode insulations; (16) cathode supports.

A section of the 6J21 is shown in Fig. 22-4. There are 12 vanes. Tuning is accomplished by both *L*- and *C*-ring plungers. A capacitive-aperture output coupling is used between one of the cavities and a horn terminated by an elliptical glass seal through which power is radiated

into the waveguide. The impedance-matching system is shown in Fig. 20-16c and is discussed in Sec. 20-10. Further details of the construction can be seen in Fig. 22-4. The loaded Q varies between approximately 50 and 150 over the tuning range. The impedance of the aperture, 2 or 3 ohms, is matched to the impedance of the waveguide at the window seal by means of a pair of exponentially tapered ramps. The seal is elliptical in shape for greater mechanical strength. Earlier rectangular windows were subjected to uneven strains when heated and quite frequently cracked under the strain. The window is usually air-cooled, but the elliptical shape has been used for hours without cooling, with no adverse effects. Early window failure in the form of a small hole punctured in the center of the window was attributed to heat caused by electron bombardment of the glass. To prevent this, iron "shoes" that extend over the top and bottom of the horn are provided with each tube. These "shoes" divert a small part of the magnetic field into the horn, thus deflecting escaping electrons so that they will not reach the window.

The thoriated-tungsten cathode is helical and bifilar. An external electromagnet supplies a magnetic field of 2200 to 3450 gauss. Operation at a plate voltage of 4000 to 5000 volts and a plate current of 300 to 400 ma is normal. A more complete summary of operating values is given in Table 22-2.

TABLE 22-2—OPERATING DATA OF MODEL 6J21 MAGNETRON

Initial filament voltage	6 volts
Initial filament current	14 5-17 5 amp
Frequency range	2460-3610 Mc

Typical Operation

Frequency	2500 Mc	3000	3500
Filament voltage (for constant filament resistance)	4 0 volts	4 3	4 3
Magnetic field*	3200 gauss	2800	2400
Anode voltage	5000 volts	5000	5000
Anode current	0 3 amp	0 30	0 30
Power output	780 watts	980	825
Efficiency	52%	65	55

* Approximate

The construction of the tuning mechanism is largely dictated by mechanical considerations. The L - and C -rings are firmly supported and guided by 12 rods. Each rod is rigidly fastened to the L - and C -rings and passes through one of the 12 cavities. Six of these are of tungsten and act only as guiding rods passing through copper bearings set in the soft-iron pole pieces built into the tube. The other six serve the same purpose below the L -ring, acting as guiding rods through bearings in the lower pole piece. Above the L -ring, however, these six are heavy copper rods supporting the L -ring and passing through the upper pole piece

without making contact. They are supported by a bellows tuning mechanism that is operated by a threaded tuning nut outside the vacuum. The copper supporting rods have large cross-sectional areas primarily to furnish a heat-conducting path to the water-cooled supporting block in the tuning mechanism.

The tuner construction is mechanically simpler if the L - and C -rings are not grounded at the periphery. However, without this grounding, considerable difficulty with extra resonances has been experienced. In the final design of the 6J21, the L -ring is grounded at the periphery by small coil springs mounted on its circumference. These make sliding contact between the L -ring and the body of the magnetron. The C -ring is not provided with r-f ground connections.

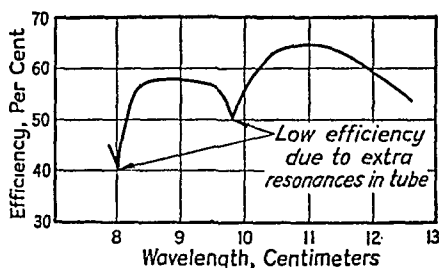


FIG. 22-5.—Efficiency curve showing spurious resonances.

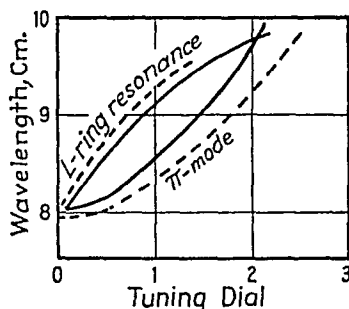


FIG. 22-6.— L -ring resonance measured by cold tests. Solid lines, L -ring slots 0.161 in. deep; dashed lines, slots 0.190 in. deep.

22-4. The Elimination of Extra Resonances.—A brief discussion of some of these extra resonances will help to point out possible techniques for eliminating them. The resonances are indicated by low-efficiency points in the tuning range and can be studied more thoroughly by cold-test methods. An efficiency curve showing two resonances at 8.1 and 9.7 cm is shown in Fig. 22-5. From the results of tests on subassemblies, it was concluded that these resonances were caused by a wave propagated around the teeth of the L -ring in such a way that its magnetic field did not pass through the bore of the L -ring. The vanes acted as tuning plungers for the resonance, giving the tuning curve shown in Fig. 22-6. By lengthening the L -ring plungers, the resonance frequency was changed so that the lower intersection with the π -mode tuning curve was shifted to 7.95 cm, *i.e.*, well outside the operating range.

The mode did not appear in grounded- L -ring tubes, and it was finally necessary to ground the L -ring to eliminate the low-efficiency point caused by the crossing at 9.7 cm.

Another resonance at 8.6 cm was found to be caused by a mode propagated around the teeth of the C -ring. Again the vanes acted as a

tuning plunger, giving a tuning curve shown in Fig. 22-7a. In cold-test studies two resonances were found, one of which did not cause trouble by intersecting with the π -mode tuning curve. The tuning curves for the π -mode and for the undesired resonances were moved relative to one another by changing the L -ring-to- C -ring spacing. The effect is shown in the transition from Fig. 22-7a to Fig. 22-7b. The tuning curve for the

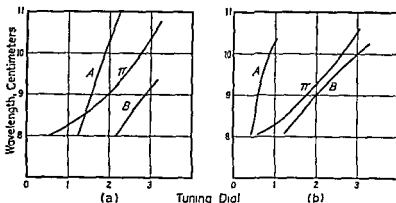


FIG. 22-7.— C -ring resonances measured by cold tests: (a) C -slots 0.167 in. deep, L -to- C clearance 0.147 in.; (b) C -slots 0.220 in. deep, L -to- C clearance 0.094 in.

π -mode is made to lie in between the extra resonances over the tuning range.

The wide tuning range is obtained necessarily by sacrifice and compromise of scaling requirements. Efficiency, mode-jump current, and extra resonances are closely tied together. When one difficulty is eliminated, another may appear. The requirements set by the applications of the magnetron, e.g., modulation band width, further complicate the picture.

22-5. The Loaded Q as a Measure of Performance.—In the development of a magnetron the efficiency, mode-jump current, and the requirements dictated by its use must be related to a common measurable factor. In this case we use the loaded Q of the magnetron. A qualitative picture of the limitations imposed may be obtained from Fig.

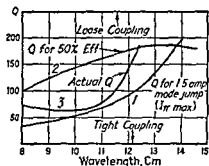


FIG. 22-8.—Limits for loaded Q .

22-8. These curves are drawn for a given internal scaling of the magnetron so that Q can be affected only by loading. Curve 1 represents the minimum Q the tube may have and meet a mode-jump current requirement of $I_{\pi \max} = 1.5$ amp. Here $I_{\pi \max}$ is the maximum value of plate current that can be reached in the π -mode before the magnetron jumps to

operation in another mode. For values of Q below curve 1 the mode jump will occur below 1.5 amp.

Curve 2 represents the *maximum* Q the magnetron may have and meet the requirement that the efficiency shall be over 50 per cent.

Curve 3 represents a typical Q for the 6J21 magnetron. This curve stays within the limits over the desired tuning range from $\lambda = 8.3$ to $\lambda = 12.2$ cm. The small distance between curves 1 and 2, particularly at the ends of the tuning range, means that even a relatively small change in the load can bring the mode-jump current below 1.5 amp or the efficiency below 50 per cent at some point in the tuning range. The center of the range is critical for mode-jump current, the ends of the range for efficiency. At wavelengths greater than the wavelength for the intersection of curves 1 and 2, it would be impossible to meet both requirements. This intersection is affected by the scaling. For the application of the tube, it is of interest to know that the operation depends very much on the load impedance. In other words, the area in the Rieke diagram where mode-jump current and efficiency are satisfactory is rather limited at some points in the tuning range.

The limitations described above make necessary very close tolerances in the spacing between the ramps in the output coupling where they join the aperture in the wall of the magnetron cavity. Small adjustments may be made by mechanically squeezing or pulling the ramps, but in general, once a tube has been sealed off, it must be very close to meeting the requirements or it will never meet them over the whole tuning range.

Knowledge of limitations placed upon the value of the loaded Q and of methods of regulating this value are useful in an understanding of the processes involved in obtaining optimum magnetron performance. Methods for determining Q exactly, however, are involved and at times uncertain.

22-6. Frequency Pushing.—In studies of modulation of the 6J21, the more easily measured quantity, frequency pushing (see Sec. 18-4), has been set up as a basis for control of band width. Frequency pushing has been defined in Sec. 21-4 as the variation of frequency of oscillation with magnetron anode current at constant load. The appearance of the output spectrum is largely determined by the shape of the frequency-pushing characteristics.

Frequency-pushing characteristics for the 6J21 are shown in Fig. 22-9. The notable feature of these curves is the large pushing obtained between the plate-current limits of 0 and 0.2 amp. This means that for 100 per cent amplitude modulation very wide frequency-modulation band widths are obtained at relatively low powers. If the amplitude modulation is restricted to 80 per cent (for 1-amp peak plate current), the frequency-modulation band width is reduced to approximately 33 per

cent of its original value. In some cases, further reduction of the amplitude-modulation percentage to values in the neighborhood of 50 per cent will almost eliminate the frequency modulation.

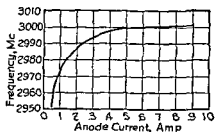


FIG. 22-9.—Typical pushing curve for 6J21 magnetron.

In order to ensure that the low-current high-pushing region is not reached, a diode "clipper" circuit may be used on the output of the modulator to limit the negative swings to voltages producing minimum currents of a specified value. It is also useful to prevent excessive upward swings in plate current by use of a "top clip-

per." Excessive upward swing has been found to reduce filament life, mode-jump current, and efficiency. Reasons for this are not thoroughly understood.

Pushing is a function of the tuning of the magnetron, as is shown in Fig. 22-10. This graph is a plot of "pushing figures" over the tuning range of a typical 6J21. These figures are selected as the variation in frequency of oscillation between the plate-current limits of 0.25 and 1.0 amp. In addition to local variations, a striking contrast is to be noted between the values of pushing at 12 cm and at 8.5 cm.

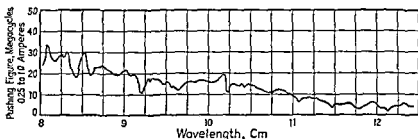


FIG. 22-10.—Variation of pushing figure over the tuning range.

22-7. External Control of Band Width.—The variation of pushing with tuning plus the necessary compromises to be made between efficiency and mode-jump limitations make some external means of controlling band width desirable. One approach to this problem is the use of an adjustable reflector in the waveguide transmission line carrying power from the tube. This reflector is placed 25 or 30 waveguide wavelengths from the magnetron output window. The resulting high- Q circuit in the output of the magnetron acts to reduce the pushing figure. The Q of the circuit increases with the length of the line between output and reflector and with the standing-wave ratio produced by the reflector. The standing-wave ratio that can be used is limited by the maximum

losses permissible in the output window of the magnetron and in the magnetron itself. The length of line is limited by the long-line effect described in Sec. 21-8. As the length of line between reflector and output window increases, the separation between multiple resonances due to the long-line effect decreases. Tuning to the desired resonance is accomplished with a line stretcher inserted in the line between magnetron and reflector.

22-8. Line Stretcher and Adjustable Reflector.—A line stretcher and an adjustable reflector suitable for use over the wide tuning range covered

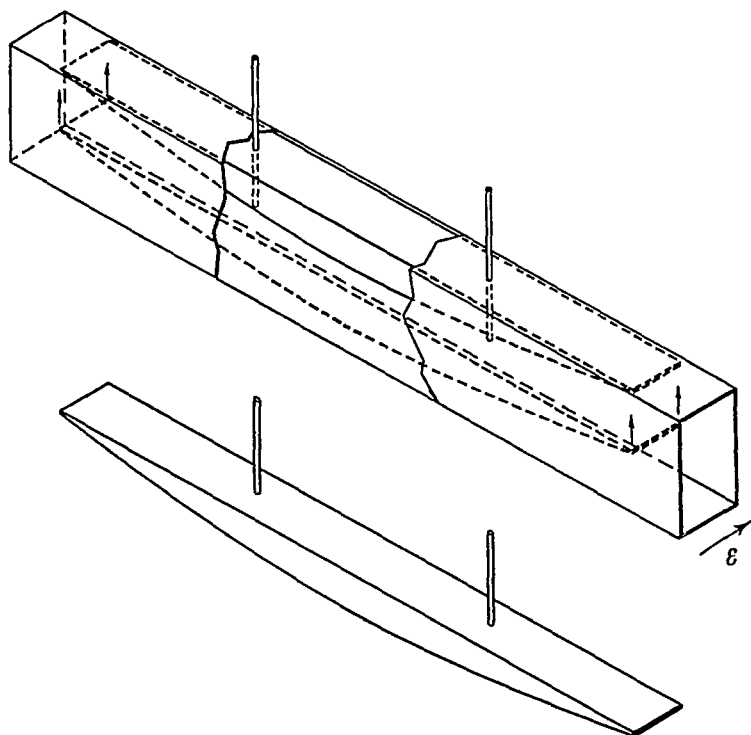


FIG. 22-11.—Dielectric-slug line stretcher.

by the 6J21 are illustrated in Figs. 22-11 and 22-12, respectively. The line stretcher is designed to provide a range in phase adjustment greater than 180 deg anywhere in the tuning range and to produce a minimum standing-wave ratio in the guide. In the model shown the minimum range is 270 electrical degrees, and the maximum standing-wave ratio produced is 1.15:1 between $\lambda = 8$ cm and $\lambda = 12$ cm. The variation in electrical length of the line is attained by moving a tapered polystyrene slug from the side of the guide, where the electric field strength is a minimum, toward the center of the guide, where the field is a maximum. To provide a greater range of phase adjustment and smaller standing-

wave ratio the slug is tapered in the \mathcal{K} dimension rather than in the \mathcal{E} dimension. Only one surface is tapered. The opposite surface is flat, so that when it is against the edge of the guide it is out of the electric field as far as possible. The metal rods supporting the slug run perpendicular to the electric field and do not introduce a noticeable reflection.

The adjustable reflector (Fig. 22-12) is designed to produce a voltage standing-wave ratio variable between the limits 1:1 and 6:1. The calibration of the adjustment can be made substantially independent of wavelength over the range of the 6J21 by proper choice of the angle that the axis of the reflecting vane makes with the \mathcal{K} dimension of the guide. For a critical angle, around 15 deg, the change with the wavelength of the

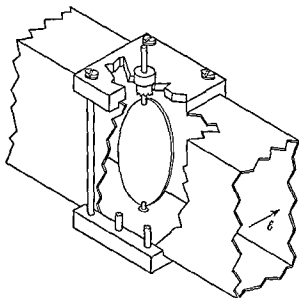


FIG. 22-12—Reflector vane.

lumped capacitive reactance offered by the vane is compensated by change in phase between reflections from different parts of the vane, so that the total effect is practically independent of the wavelength between 8 and 12 cm. The axis of rotation of the vane is perpendicular to the electric field in the waveguide. This eliminates the need for good electrical contact where the vane axle goes through the side of the guide. When the plane of the vane is parallel to the electric field, a maximum reflection is produced; when the plane of the vane is perpendicular to the electric field, a minimum reflection is produced. The minimum standing-wave ratio is less than 1.05:1 over the range 8 to 12.5 cm.

22-9. Filament Emission and Back-heating.—The filament of the 6J21 is of thoriated tungsten, carburized to improve thoria distribution. This type of cathode is more rugged than oxide-coated cathodes but more

constant. For some types of experimental operation a manually controlled bridge-balancing circuit may be used. However, for most work automatic controls are much to be preferred. Three general types have been devised. In one, a motor-operated control, the unbalance voltage from a bridge in the filament circuit is applied to a two-phase motor and turns the variac that controls the filament voltage. The bridge balances filament voltage against filament current. The control keeps the ratio between these quantities constant. This circuit is shown in Fig. 22-13.

Two other types of circuit have been used. One completely electronic circuit utilized an oscillator and an amplifier as the power source for the filament transformer. The feedback to the oscillator is the difference between a positive feedback voltage proportional to filament current and a negative feedback voltage proportional to filament voltage. Any change in filament resistance is compensated by the change in oscillator output. This method has the disadvantage of being bulky and the advantage of being quick-acting.

A third type of circuit, using a saturable reactor in series with the filament transformer, requires fewer components and gives speedy action. The difference of two rectified voltages proportional, respectively, to filament voltage and filament current is applied to the grid of a receiving tube feeding the saturable reactor. Any unbalance will change the impedance of the saturable reactor in such a way as to restore balance.

Another fact that must be taken into account when helical filaments are used is the effect of filament current on magnetic field. A ripple in the magnetic field caused by alternating current in the filament shows up as amplitude and frequency modulation in the output. From the variation in anode current that is easily observed on an oscilloscope, and knowledge of the d-c operating point, the amount of frequency modulation is easily calculated from a pushing characteristic. It can also be measured directly by means of a spectrum analyzer. Monofilar filaments in early models of the 6J21 caused 8 to 12 Mc of frequency modulation. Use of bifilar filaments reduces this to 1 to 4 Mc. The effect can of course be entirely eliminated by using d-c heating of the filament, but this is not very practical where the filament is at high voltage. Compensation for back-heating tends to reduce filament modulation of the magnetic field by reducing the a-c heating current.

22-10. The Operating Characteristics of the 6J21.—In order to give a complete picture of the operating characteristics of the 6J21 tube, a number of graphs should be included here such as performance charts and Rieke diagrams for two or three points in the tuning range; but such diagrams are not yet available for the tube in its latest stage of development.

The variation of efficiency and mode-jump current with wavelength is

shown in Fig. 22-14. This graph and the pushing diagram, Fig. 22-10, give three of the more important characteristics of the 6J21 magnetron. Another characteristic of interest in connection with the use of the tube is the modulation impedance as obtained from the slope of the curves for constant magnetic field in the performance chart; it varies from about 500 ohms at 8 cm to about 1500 ohms at 12 cm.

The performance of the tube at high standing-wave ratio is characterized by substantially reduced efficiency. For a standing-wave ratio of 3:1 in the low-conductance region of the Rieke diagram, the efficiency is 30 to 40 per cent, which compares unfavorably with 50 to 60 per cent at unity standing-wave ratio. The higher standing-wave ratio also causes a 20 to 25 per cent increase in back-heating. Increased stability in operation of the 6J21 has been obtained by using a series magnet, as described in Sec. 21-5. Over the wide tuning range the tendency to operate at higher anode currents at the long wavelength end is checked by the greater magnetic field produced. It is recommended that both series-magnet and automatic back-heating control be used for increased tube life.

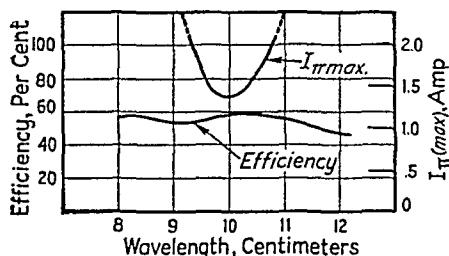


FIG. 22-14.—Variation of mode-jump current and efficiency over the tuning range.

22-11. The 4J60-to-4J65 Low-power Magnetrons.—The 50-to-100-watt tubes for frequencies in the neighborhood of 3000 Mc were part of a coordinated large-scale program of tube and equipment development that reached quantity production before the end of the war. The equipment in which these tubes were used required that the tube could be modulated, that the output should be maintained at a reasonable level in spite of large fluctuations in supply voltage, that the cooling system for the anode should be as simple and light as possible, and that the frequency of oscillation should be adjustable by remote control.

The solution of these requirements is a compact little tube with forced-air cooling and a permanent magnet integrated with the tube. The voltage variations are taken care of in the transmitter by connecting in series with the magnetron a pentode or beam-power tube, which functions both as current regulator and as output stage in the modulator. The high-impedance or constant-current characteristic of the modulator tube reduces the effect of the line-voltage fluctuations very much. This tube can be given a third function, *i.e.*, to act as a limiter to prevent the anode current from reaching the mode-jump current at the high-current end and from dropping down into the low-power operating region where very severe pushing occurs.

The 4J60-to-65 tubes are 18-vane double-ring-strapped magnetrons with the "cookie-cutter" variety of *C*-ring tuner (Sec. 20-11). Even though very small clearances are used between *C*-ring and straps, the tuning range is only about 12 per cent. Table 22-3 shows the tuning ranges for the various tubes in this group.

TABLE 22-3.—TUNING RANGES OF THE 4J60-TO-65 MAGNETRONS

4J60	2230-2470 Mc
4J61	2450-2720 Mc
4J62	2700-3010 Mc
4J63	2990-3330 Mc
4J64	3310-3670 Mc
4J65	3650-4030 Mc

The output coupling is a simple loop in one of the anode cavities. The size of the loop is adjusted to give the best compromise between modula-

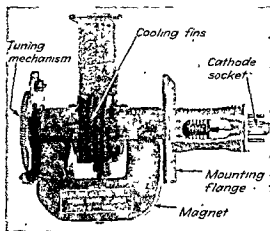


FIG. 22-15.—The 4J60 magnetron.

tion band width and efficiency, on one side, and mode-jump current on the other. The coupling loop is directly connected to the coaxial transmission line without any matching network. It has been found advisable, however, to have a line stretcher in the line close to the output terminals so that the holes in the tuning range due to the long-line effect can be adjusted.

A typical tube is shown in Fig. 22-15. The resonant system and the electrodes are all inside the small central body, which is provided with cooling fins; the bulk consists of auxiliary parts, including magnet, drive for the tuner, cathode socket, output terminals, etc. The emitter is an indirectly heated oxide-coated nickel-mesh cathode.

Operating data are given in Table 22-4.

TABLE 22-4.—OPERATING DATA FOR 4J60 MAGNETRON
Oxide-coated Cathode

Heater voltage.....	6.3 volts
Heater current.....	3.8 amp
Frequency range.....	2230-2470 Mc

Typical Operation

Heater voltage.....	6.3 volts
Anode voltage.....	1200 volts
Anode current.....	0.125 amp
Power output.....	65 watts
Efficiency.....	43%

Maximum Ratings

Peak anode current.....	0.275 amp max
Average anode current.....	0.150 amp max
Anode dissipation.....	100 watts max
Anode temperature.....	100°C max
Cooling, forced air	

The obvious disadvantage of these tubes is the narrow tuning range. Six tubes are required to cover a frequency range of less than 2 to 1. Shift of frequency in a transmitter requires frequent replacement of tubes, an operation that is inconvenient even with the most streamlined plug-in units. Cost and bulk are also considerable; in fact, would soon become prohibitive if a larger frequency range were required. The fact that these tubes are available today is explained by the situation at the time they were developed. The early multicavity magnetrons were spot-frequency oscillators, which sufficed for radar applications. When the need was felt for low-power tunable c-w magnetrons, it was realized that the time required for development increased at least as fast if not faster than the tuning range. The need was so urgent that only a small tuning range was requested. The first experimental tubes were actually built for only 6 per cent range.

The successful development of the 6J21 demonstrates what can be done with combined *L*- and *C*-rings. A tuning range of 1.5:1 is a considerable improvement. If the 6J21 anode and tuner were to be scaled down to 1200 volts, the dimensions and clearances would become so small as to present a very difficult manufacturing problem. It is by no means hopeless, however, to attempt to use the same general idea, modified to suit the smaller tubes.

22-12. The Squirrel-cage Magnetron or Donutron.—A different approach to the multisegment magnetron is shown in Fig. 22-16. The resonator is a short cylindrical cavity, and the anode segments extend as

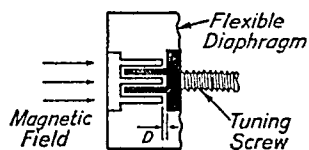


FIG. 22-16.—Schematic representation of a donutron (cathode assembly omitted). One anode half has its fingers shaded and the other drawn in outline for the sake of clarity.

fingers from the two flat sides of the cavity. In the simplest form of this tube, as shown in the figure, every other finger is attached to the same side of the cavity. The fingers thus form a "squirrel cage" around the cathode, and the cavity outside the electrode system is shaped somewhat like a doughnut. One of the flat sides of the cavity is flexible, tuning being accomplished by moving the two finger systems axially, relative to each other. Roughly we can consider the finger systems as a capacitive loading on the cavity resonator. Variation of the capacitance changes the resonance frequency of the cavity.

22-13. Resonance Modes.—If, as a first approximation, the electromagnetic-field configuration is assumed to be the same in all planes perpendicular to the axis of the electrode system, the space analysis of the field into modes and component waves that revolve around the cathode will be exactly the same as for the multicavity magnetron (Sec. 20-5). The resonant system that determines the frequencies of the various modes is quite different, however, and the discussion of the operation, advantages, and disadvantages of the donutron will necessarily center around the modes and mode separation of the resonant system.

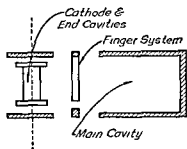


FIG. 22-17.—The three components of the resonant system.

It is convenient to think of the resonant system as a combination of three separate components, the annular cavity, the finger systems, and the interaction space plus end cavities (Fig. 22-17). For all the resonance modes of interest here, the dimensions of the interaction space are small compared with the wavelength. Consequently, the interaction space is nonresonant, contributing only a pure susceptance to the equivalent network of the resonator. The finger system can act either as a pure capacitance, when the instantaneous voltage between the two halves is the same at all points, or as a resonant system, when a standing wave exists in the system, the zigzag slot between the halves acting as a reentrant transmission line, as illustrated in Fig. 22-18. If the zigzag shape is disregarded and the slot is considered to be a smooth line, the transverse voltage varies with the azimuth coordinate θ , as indicated by the relation

$$V_r = A_p e^{i\omega t} \sin p\theta + B_p e^{i\omega t} \cos p\theta \quad (22-1)$$

where A_p and B_p are constants to be determined from the boundary conditions, and the resonance modes of the line are found by giving p all integral values from 0 to ∞ . The resonance frequencies of these modes depend on the length of the fingers and on the spacing and radius of the finger system.

The resonance modes of the annular cavity also give an axial voltage difference of the form of Eq. (22-1). The radial variation of this voltage is

$$V_r = P_p e^{j\omega t} J_p(kr) + Q_p e^{j\omega t} N_p(kr) \quad (22-2)$$

where J_p and N_p represent the Bessel and Neumann functions of order p , k is $2\pi/\lambda$, r the radius, and P_p and Q_p are constants to be determined by the boundary conditions. The resonance frequencies of the various modes are determined by the maximum and minimum radii of the cavity.

To find the resonances of the whole system from theory, the susceptance of the three parts of the system, as seen from the boundary lines and surfaces between them, should be calculated. Plotting these susceptances as functions of frequency, adding them, and finding the frequencies for which their sum is zero would give the resonance frequencies of the donutron. This procedure must be carried out for all values of p in order to find all the resonance modes. As the interaction space can be considered to be nonresonant and the reentrant line has only one resonance for each value of p , this procedure will be exactly the same as that used to analyze the long-line effect in Sec. 21-8. As in that analysis, the zeros that represent admittance maximums, or series resonance, should be disregarded. The resonance frequencies are illustrated in Fig. 22-19.

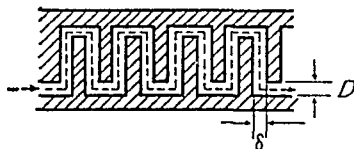


FIG. 22-18.—Reentrant line formed by finger system.

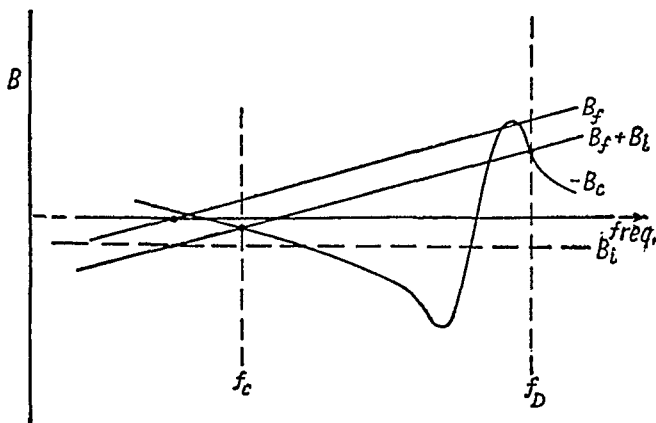


FIG. 22-19.—Resonance frequencies of the donutron (modes C and D). B_i , susceptance of interaction space, B_f , susceptance of finger system; B_c , susceptance of cavity; f_c and f_d , frequencies of resonance modes C and D .

22-14. Resonance Modes from Cold Tests.—The resonance modes have been studied by cold tests by means of the test equipment indicated schematically in Fig. 22-20. An oscillator is modulated 100 per cent by a square wave of 2000-cycle frequency, and its output is fed through a

lossy line to the output loop of the magnetron to be tested. The cathode of this tube has been replaced by a dummy that is provided with a radial probe protruding a small distance outside the cylindrical surface of the cathode. This dummy cathode is kept revolving by a synchronous motor.

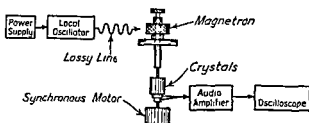


FIG. 22-20.—Cold-test equipment.

motor. The modulated r-f power picked up by the probe is fed to a crystal detector, which is connected to an audio amplifier. The output of this amplifier is applied to the vertical plates of a cathode-ray oscilloscope, the linear time sweep of which is synchronized with the motor that drives the probe. The screen of the oscilloscope will consequently show a pattern that is a plot of the square of the radial electric field at the cathode vs. the azimuthal coordinate θ . The tangential component is proportional to this radial component but 90 deg out of phase in space. This fact ties these patterns in with the analysis of the tangential component of the electric field given in Sec. 20-5 in terms of pairs of waves traveling around the interaction space with one wave traveling in each direction. Now of course any pair of components can be expressed as the sum either of two traveling or of two stationary components according to the identity

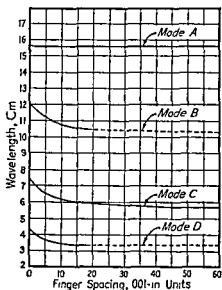


FIG. 22-21.—Resonance modes of the donutron (tube 39).

two stationary components according to the identity

$$Ae^{j(\omega t + \theta)} + Be^{j(\omega t - \theta)} = Pe^{j\omega t} \sin \theta + Qe^{j\omega t} \cos \theta \quad (22-3)$$

where A , B , P , and Q are constants determined by the boundary conditions.

As the resonance modes are made up of component waves of this kind, the same identity applies to the modes. As pointed out in Sec. 20-5, the π -mode has always a standing-wave nature, because there are equal com-

ponents of opposite rotation of all the waves comprising this mode. All other modes will also appear in pairs, of which both members can be

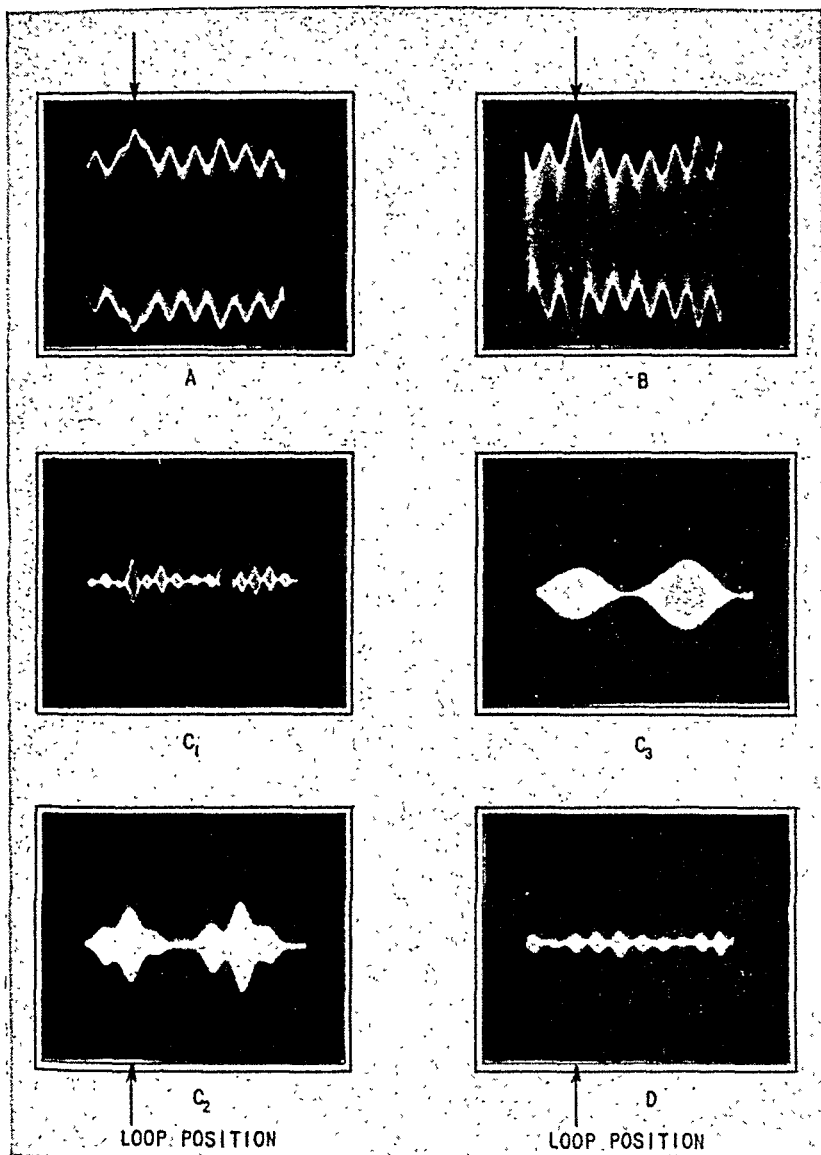


FIG. 22-22.—Cold-test patterns (tube 39, 16 fingers).

thought of as identical sets of component waves, different only in the direction of rotation, or as two standing-wave systems displaced 90 deg from each other in space.¹ In considering the electron interaction, the

¹ See footnote 1, p. 484.

concept of traveling waves is helpful; the analysis of the resonance modes of the system, however, is easier in terms of the stationary wave patterns.

Figures 22-21 and 22-22 show the tuning curves and cold-test patterns for a donutron with 16 equal fingers. Four principal types of oscillations are indicated. The modes designated *A* and *B* in the figures both give cold-test oscillograms of the form $(a \pm b \sin \theta)^2$, the only difference between the two being the sign between the two terms. These patterns represent a π -mode oscillation that is unbalanced with respect to the cathode. The main reason why *A* and *B* are considered different modes is the fact that mode *A* is practically unaffected by tuning, while mode *B* is tunable. The explanation probably is that in mode *A* the circuit between anode and cathode (the "unbalanced" circuit) determines the frequency of resonance, while in mode *B* the fundamental mode of the cavity and also the capacitance of the finger system determine the frequency. Mode *A* gives small voltage between the fingers and has never produced much power in hot operation.

If we disregard the unbalance, mode *B* represents the operation for which the finger magnetrons originally were designed. In this mode the r-f electric field outside the finger system is parallel to the axis of the tube and independent of the azimuthal coordinate θ ; it falls off to zero at the circumference of the cavity. The r-f magnetic lines outside the fingers are circles concentric with the tube axis. Operation with efficiencies up to 35 per cent has been obtained in this mode, but the output is seriously limited by a low mode-jump current. The main cause of this probably is that the heavy capacitive loading that the finger system represents makes $\sqrt{L/C}$ and the shunt impedance low. Because mode-*C* operation appeared more promising, the possibilities of mode-*B* operation have not been thoroughly investigated. Tuning ranges of 1.1 to 1 to 1.3 to 1 have been obtained with plain fingers; ranges up to 1.7 to 1 or more have been recorded for tubes in which the capacitance at the tips of the fingers has been increased by tabs or rings fastened to the finger tips.

22-15. C-mode Operation.—Mode *C* is the resonance of lowest frequency for $p = 1$ in Eqs. (22-1) and (22-2) and corresponds to a standing wave of 1 wavelength on the reentrant line formed by the finger system. The r-f electric field in the cavity is axial in this mode also, and its maximums coincide in azimuth with the voltage maximums on the finger system; and the voltage nodes and the zeros of the electric field in the cavity occur at the same azimuth. The shape of the lines of the r-f magnetic field is indicated in Fig. 22-23. The radial magnetic field is zero at the voltage antinode and maximum at the voltage node on the finger system. The coupling loop is shown in the position of maximum coupling, i.e., displaced 90 deg from the nodal plane of the finger system.

Because the admittance of the reentrant line formed by the finger

system is considerably larger than the admittance of the cavity, the resonance frequency is largely determined by the line. This makes it possible to predict the wavelength from the length of the zigzag line formed by the gap between the fingers. When the tuner is set so that this gap is constant at all points of the line, it has been found that for all known finger magnetrons the wavelength is approximately *half* the total length of the zigzag line, *i.e.*,

$$\lambda \approx \frac{2\pi r_a + Nl}{2} \quad (22-4)$$

where r_a is the anode radius, N the number of fingers, and l the length of the fingers. Changing the dimensions of the cavity has been found to have only a relatively small influence on the frequency of resonance.

It should be emphasized that this is a purely empirical formula. If the line were unfolded, the first resonance should of course occur for a wavelength equal to the full length of the line. The folding of the line reduces the distributed inductance, because a large part of the current will by-pass the fingers, and it happens that for practical dimensions this effect will cut the wavelength approximately in half.

The cold-test oscillograms C_1 , C_2 , and C_3 in Fig. 22-22 indicate that the radial electric field close to the cathode varies with the azimuth according to the relation

$$\varepsilon_r = \left(A \sin \theta \sin \frac{N}{2} \theta + B \sin \theta \right) e^{j\omega t} \quad (22-5)$$

A and B vary with the axial position of the probe along the cathode. At the mid-point of the cathode, B would be expected to be zero. This is not quite true in the scope picture C_1 , which indicates that the geometrical and electrical symmetry of the system was not perfect under the conditions for which this picture was taken. Beyond the finger tips, close to the wall of the end cavity, A would be expected to be zero; the pattern C_3 approximates $(\sin \theta)^2$ and confirms this expectation.

For the electron interaction in the tube the first term in Eq. (22-5) is the important one. It can be expanded as follows:

$$\begin{aligned} \varepsilon_r = A e^{j\omega t} \sin \theta \sin \frac{N}{2} \theta &= \frac{1}{4} A \left(e^{j[\omega t + (\frac{N}{2}-1)\theta]} + e^{j[\omega t - (\frac{N}{2}-1)\theta]} \right. \\ &\quad \left. - e^{j[\omega t + (\frac{N}{2}+1)\theta]} - e^{j[\omega t - (\frac{N}{2}+1)\theta]} \right) \quad (22-6) \end{aligned}$$

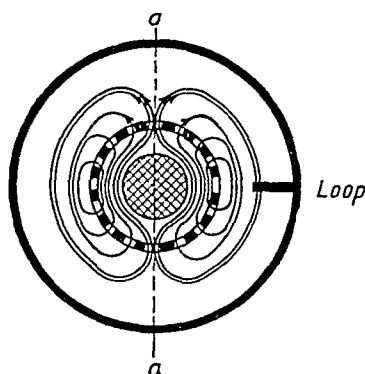


FIG. 22-23.—R-f magnetic field of mode C.

Instead of the π -mode, for which $n = N/2$, we thus have in mode- C operation, in the interaction space, two pairs of revolving modes with mode numbers $n = N/2 + 1$ and $N/2 - 1$, respectively. In the mode-line diagram of Fig. 20-15 the two lines representing operation in which the electrons synchronize with these components are fairly close together if N is large compared with unity; because the frequency of resonance is the same, the angular velocity of each component is inversely proportional to its mode number. The operation with $n = N/2 + 1$ is most likely to be established because this mode line is lower and will be reached first when the anode voltage is building up.

22-16. The Phase-reversing Anode.—An improvement in performance of C -mode operation has been obtained with the *phase-reversing anode*. The construction of such an anode and a comparison between the instantaneous voltages on the fingers of a regular anode and a phase-reversing anode are shown in Fig. 22-24, where the finger systems are shown unfolded. On the symmetrical anode A in C -mode operation, two adjacent fingers at each nodal plane in the cavity will have voltages of the same polarity. The revolving space-charge cloud will be subject to a more uniform field configuration if the polarity of adjacent fingers alternates regularly all the way around the anode. This has been accomplished in the phase-reversing anode simply by a slight rearrangement of the fingers, as shown in B Fig. 22-24. The sine curves represent the potential distribution on the walls of the cavity.

FIG. 22-24.—Comparison of symmetrical anode A and phase-reversing anode B .

Each finger is attached to the side that gives the finger the desired polarity. At each voltage node two adjacent fingers are attached to the same side of the cavity and can be made as one broad finger. The radial r-f electric field component now has the approximate value

$$\begin{aligned} E_r &\approx \frac{2}{\pi} A e^{j\omega t} \sin \frac{N}{2} \theta \left(1 + \frac{2}{3} \cos 2\theta - \frac{2}{15} \cos 4\theta + \dots \right) \\ &\approx \frac{1}{\pi} A \left[e^{j(\omega t + \frac{N}{2}\theta)} - e^{j(\omega t - \frac{N}{2}\theta)} \right] + \frac{1}{3\pi} A \left\{ e^{j[\omega t + (\frac{N}{2}-2)\theta]} - e^{j[\omega t - (\frac{N}{2}-2)\theta]} \right\} \\ &\quad + \frac{1}{3\pi} A \left\{ e^{j[\omega t + (\frac{N}{2}+2)\theta]} - e^{j[\omega t - (\frac{N}{2}+2)\theta]} \right\} + \dots \quad (22-7) \end{aligned}$$

The fundamental component has increased in amplitude in the ratio $4:\pi$, and only one component and one mode line must be considered, because the other terms in the expansion are of appreciably smaller amplitude. We can say that the phase-reversing anode makes possible

the operation of the electronic system in the π -mode ($n = N/2$) in spite of the fact that the cavity and the reentrant line oscillate in the mode for which $p = 1$. A third advantage of this anode design is that oscillation in the *A* and *B* modes is just as much discouraged as the *C* mode is encouraged. Actually, the *A* and *B* modes have not been observed at all in tubes with phase-reversing anodes.

22-17. Mode Degeneracy.—The degeneracy of the *C* mode has presented some problems in finger magnetrons. As already mentioned each mode except the π -mode is actually a degenerate pair and we can think of each pair as two field patterns, the nodal lines and maximums of which are 90 deg apart in space. Now if the geometry of the system is not perfectly symmetrical or, as is usually the case, a coupling loop is inserted at one point in the cavity, the degeneracy may be removed, so that the resonance frequencies of the pair may differ slightly. A certain amount of coupling may then also exist between the two modes.

Figure 22-25 shows tuning curves for a mode pair of this kind. These observations were made by cold tests on a tube with symmetric anode. The modes have been identified by their position in relation to the coupling loop; one with practically no coupling to the loop has been called the *red* mode; the other one with maximum coupling to the loop the *black* mode. Field patterns from cold tests for the red and black modes are shown in Fig. 22-26 for various positions of the probe, beginning at the mid-point of the cathode and moving toward the wall of the cavity. The fact that no pattern is quite pure indicates the presence of both modes; in the red-mode pattern higher modes must also be assumed present to explain the irregularities.

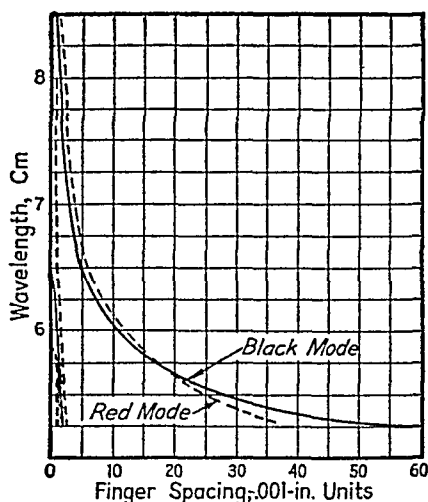
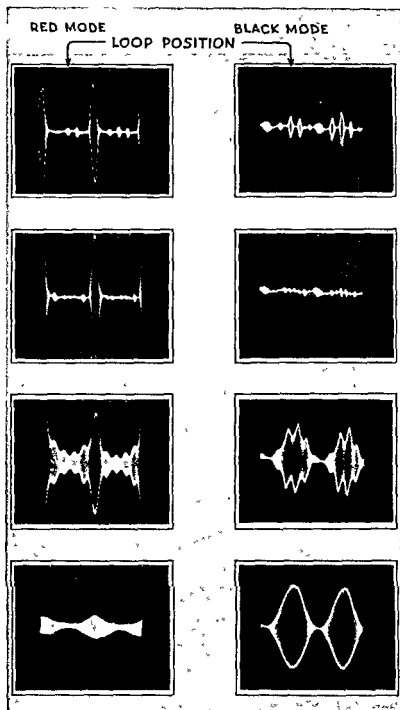


FIG. 22-25.—Tuning of black and red modes (mode *C*, tube 47).

The presence of two modes coupled to each other and of nearly but not quite equal resonance frequency leads to instability of operation and is highly undesirable. In order to change the resonance frequency of the red mode substantially without affecting the black mode, the equivalent capacitance of the mode can be increased by adding a capacitance across the line at the point where the voltage in the black mode is zero and the voltage in the red mode is a maximum. Figure 22-27 shows a way of accomplishing this. Two extra fingers are added outside the regular

FIG. 22-26.—Cold-test patterns in *C*-mode excitation.

finger system in such a way that they form capacitors with the wide fingers (obtained by combining two fingers) in a phase-reversing anode. The effect on the tuning curves and mode separation is demonstrated in Fig. 22-28. The alternate possibility of increasing the capacitance at

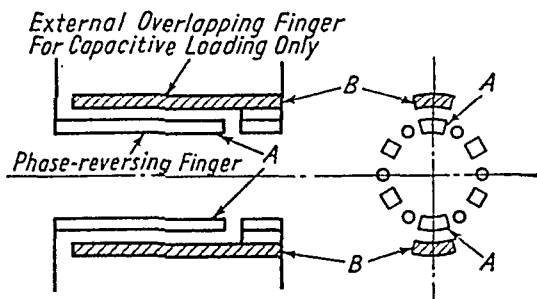


Fig. 22-27.—Longitudinal and transverse cross sections of overlap phase-reversing anode.

the tips of all fingers except the two wide fingers has also been tried successfully. The red curve then falls below the black curve.

The performance in *C*-mode operation has been considerably better than in the *B* mode. Output power of about 50 watts has been obtained consistently over appreciable periods of time, though adequate life tests have not yet been performed. Considerably higher power has been recorded on short tests. The efficiency is in the range 30 to 50 per cent.

22-18. The Effect of Finger Shape on Tuning Range.—The tuning range varies considerably with the number and length of the fingers. A large number of long fingers has in general given rather poor results, but there are indications that this is due to mechanical difficulties and lack of precision in the tuning mechanism. If the gaps at the tips of the fingers do not decrease uniformly at all fingers when the tube is tuned toward longer waves, a large number of modes appear at the long-wave end of the range differing in wavelength by only a few millimeters.

Recent results with improved tuning mechanisms have demonstrated satisfactory operation over a tuning range of more than 1.5 to 1. Figure 22-29 shows the tuning curve and output power for constant input of a tube with a phase-reversing anode and overlapping fingers for elimination of the red mode.

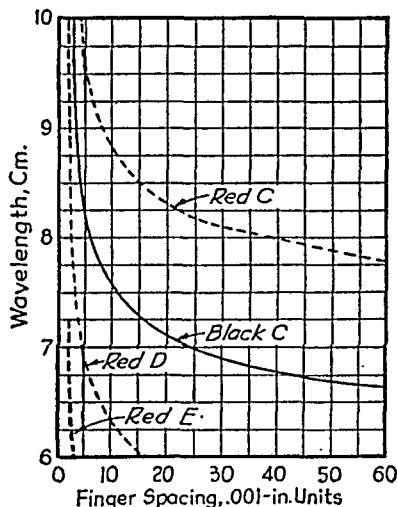


Fig. 22-28.—Cold resonances (overlap model, tube 51).

By increasing the capacitance at the finger tips with tabs, tuning ranges up to 2 to 1 have been obtained, but the operation was marred at the long-wave end of the range by the millimeter moding mentioned above. It is evident, however, that the results quoted here by no means approach the limits of what can be accomplished with these tubes. Several schemes for improving the linearity of the tuning curves have been suggested, but no experimental data are as yet available.

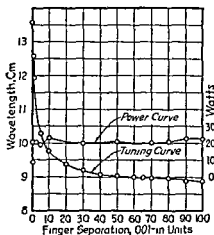


FIG. 22-29.—Tuning and output curves for overlap phase-reversing model (tube 57).

Another subject that has not yet been thoroughly investigated is the nature and possibilities of the higher modes, *D*, *E*, etc., which have been observed in cold tests. They all seem to belong to the family for which $p = 1$. They have not been found in hot operation.

In conclusion it can be stated that the donutron with the overlap phase-reversing anode shown in Fig. 22-27 has demonstrated the possibility of building a finger magnetron with appreciable tuning range, simple construction, and remarkable freedom from mode trouble. The difficulty of cooling the narrow fingers, however, seems to limit this design to low-power tubes.